



US009297028B2

(12) **United States Patent**
Donaldson et al.(10) **Patent No.:** **US 9,297,028 B2**
(45) **Date of Patent:** ***Mar. 29, 2016**(54) **FERMENTIVE PRODUCTION OF FOUR CARBON ALCOHOLS**(75) Inventors: **Gail K. Donaldson**, Newark, DE (US);
Lixuan Lisa Huang, Hockessin, DE (US); **Lori Ann Maggio-Hall**,
Wilmington, DE (US); **Vasanthanagarajan**,
Wilmington, DE (US); **Charles E. Nakamura**,
Claymont, DE (US); **Wonchul Suh**, Hockessin, DE (US)(73) Assignee: **Butamax Advanced Biofuels LLC**,
Wilmington, DE (US)(*) Notice: Subject to any disclaimer, the term of this
patent is extended or adjusted under 35
U.S.C. 154(b) by 1615 days.

This patent is subject to a terminal disclaimer.

(21) Appl. No.: **11/527,995**(22) Filed: **Sep. 27, 2006**(65) **Prior Publication Data**

US 2008/0182308 A1 Jul. 31, 2008

Related U.S. Application Data(60) Provisional application No. 60/721,677, filed on Sep.
29, 2005, provisional application No. 60/814,470,
filed on Jun. 16, 2006.(51) **Int. Cl.****C12P 7/16** (2006.01)
C12N 9/10 (2006.01)
C12N 9/04 (2006.01)
C12N 9/88 (2006.01)
C12N 9/02 (2006.01)
C12N 9/06 (2006.01)
C12N 1/20 (2006.01)
C12N 15/00 (2006.01)
C07H 21/04 (2006.01)(52) **U.S. Cl.**CPC . **C12P 7/16** (2013.01); **C12N 9/001** (2013.01);
C12N 9/0004 (2013.01); **C12N 9/0006**
(2013.01); **C12N 9/0008** (2013.01); **C12N**
9/1029 (2013.01); **C12N 9/88** (2013.01); **C12Y**
101/01157 (2013.01); **C12Y 103/99002**
(2013.01); **C12Y 203/01009** (2013.01); **C12Y**
402/01055 (2013.01); **Y02E 50/10** (2013.01);
Y02P 20/52 (2015.11)(58) **Field of Classification Search**CPC **C12N 9/1029**; **C12N 9/0006**; **C12N 9/88**;
C12N 9/001; **C12N 9/0008**; **C12P 7/16**

See application file for complete search history.

(56) **References Cited****U.S. PATENT DOCUMENTS**2,198,104 A 4/1940 Carnarius et al.
4,424,275 A 1/1984 Levy4,520,104 A 5/1985 Heady et al.
4,568,643 A 2/1986 Levy
5,063,156 A 11/1991 Glassner et al.
5,192,673 A 3/1993 Jain et al.
5,210,032 A 5/1993 Kashket
5,424,202 A 6/1995 Ingram et al.
6,358,717 B1 3/2002 Blaschek et al.
6,835,820 B2 * 12/2004 Cannon et al. 536/23.2
6,960,465 B1 11/2005 Papoutsakis et al.
8,372,612 B2 * 2/2013 Larossa et al. 435/148
8,426,173 B2 * 4/2013 Bramucci et al. 435/160
8,518,678 B2 * 8/2013 Flint et al. 435/160
8,637,289 B2 * 1/2014 Anthony et al. 435/194
2002/0028492 A1 3/2002 Lenke et al.
2002/0182690 A1 * 12/2002 Cannon et al. 435/135
2004/0234649 A1 11/2004 Lewis et al.
2007/0259411 A1 11/2007 Bramucci et al.
2008/0124774 A1 5/2008 Bramucci et al.
2008/0138870 A1 6/2008 Bramucci et al.**FOREIGN PATENT DOCUMENTS**AU 620802 2/1992
CA 2039245 3/1991
EP 0 112 459 A1 7/1984
EP 0 282 474 A1 9/1988
EP 0 315 949 A1 5/1989
EP 1 149 918 A1 4/2000
JP 61-209594 9/1986
JP 63-017695 4/1988
JP 63-102687 5/1988
JP 63-254986 10/1988
WO WO 90/02193 A1 3/1990
WO WO 98/51813 A1 11/1998
WO WO 2007/041269 A2 4/2007
WO WO 2008/006038 A2 1/2008
WO WO 2008/052991 A2 5/2008
WO WO 2008/072920 A1 6/2008
WO WO 2008/072921 A1 6/2008**OTHER PUBLICATIONS**Branden et al. Introduction to Protein Structure, Garland Publishing
Inc., New York, p. 247, 1991.*
Durre et al. Handbook on Clostridia. CRC Press, Mar. 29, 2005, 177,
182-184, and 212-220.*
Stewart et al. Biotechnology and Genetic Engineering Reviews,
14:67-143, 1997.*
Wallace et al. Purification of crotonyl-CoA reductase from
Streptomyces collinus and cloning, sequencing and expression of the
corresponding gene in *Escherichia coli*. Eur. J. Biochem. 233, 954-
962 (1995).*
Ryu et al. Recent progress in biomolecular engineering. Biotechnol
Prog. Jan.-Feb. 2000;16(1):2-16.*

(Continued)

Primary Examiner — Yong Pak

(57) **ABSTRACT**Methods for the fermentive production of four carbon alco-
hols is provided. Specifically, butanol, preferably 1-butanol is
produced by the fermentive growth of a recombinant bacte-
rium expressing a 1-butanol biosynthetic pathway.

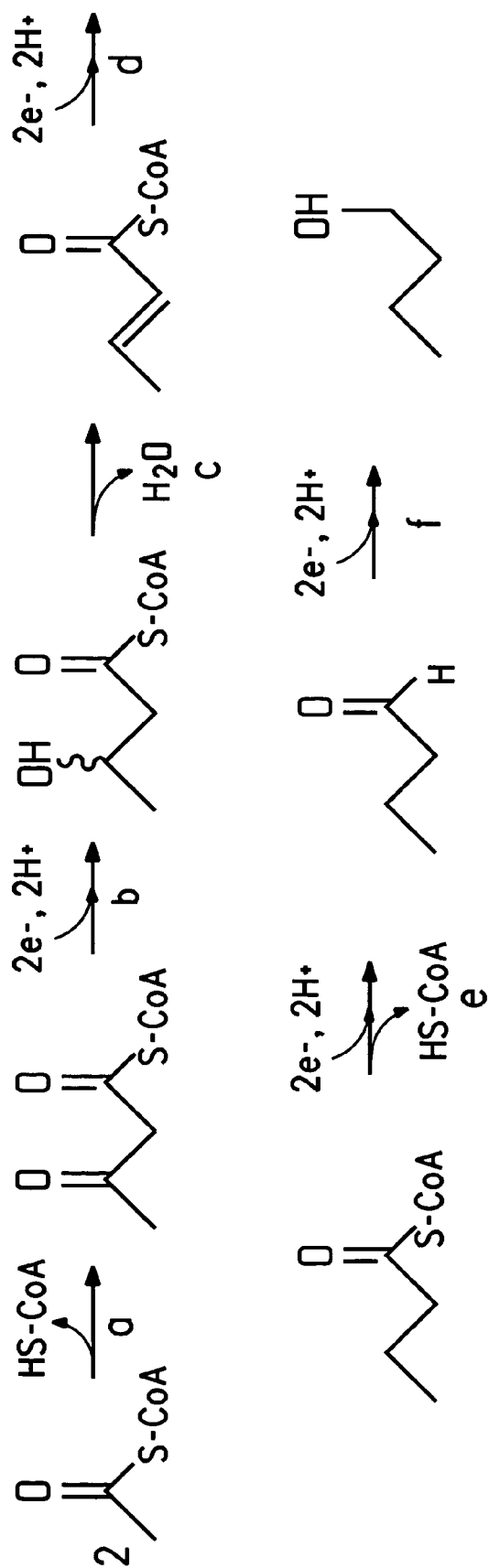
(56)

References Cited

OTHER PUBLICATIONS

- Nielsen et al. Metabolic engineering: techniques for analysis of targets for genetic manipulations. *Biotechnol Bioeng.* Apr. 20-May 5, 1998;58(2-3):125-32.*
- Deken. The Crabtree-effect: A regulatory system in yeast. *J. Gen. Microbiol.* 44, 149-156. 1966.*
- Ullmann's Encyclopedia of Industrial Chemistry, 6th Edition, 2003, vol. 5:716-719.
- Girbal et. al., Regulation of Solvent Production in *Clostridium acetobutylicum*, *Trends in Biotechnology*, 1998, vol. 16:11-16.
- Fontaine et. al., Molecular Characterization and Transcriptional Analysis of ADHE2, the Gene Encoding the NADH-Dependent Aldehyde/Alcohol Dehydrogenase Responsible for Butanol Production in Alcohologenic Cultures of *Clostridium acetobutylicum* ATCC 824, *Journal of Bacteriology*, 2002, Pages vol. 184:821-830.
- Cornillot et. al., The Genes for Butanol and Acetone Formation in *Clostridium acetobutylicum* ATCC 824 Reside on AL Large Plasmid Whose Loss Leads to Degeneration on the Strain, *Journal of Bacteriology*, 1997, vol. 179:5442-5447.
- Bermejo et. al., Expression of *Clostridium acetobutylicum* ATCC 824 Genes in *Escherichia coli* for Acetone Production and Acetate Detoxification, *Applied and Environmental*, 1998, vol. 64:1079-1085.
- D.R. Woods, The Genetic Engineering of Microbial Solvent Production, *Trends in Biotechnology*, 1995, vol. 13:259-264.
- P. Durre, New Insights and Novel Developments in Clostridial Acetone/Butanol/Isopropanol Fermentation, *Applied Microbiology and Biotechnology*. 1998, vol. 49:639-648.
- Harris et. al., Characterization of Recombinant Strains of the *Clostridium acetobutylicum* Butyrate Kinase Inactivation Mutant: Need for New Phenomenological Models for Solventogenesis and Butanol Inhibition, *Biotechnology and Bioengineering*, 2000, vol. 67:1-11.
- Crystal structure 1ZSY entry in RCSB Protein Data Bank (www.pdb.org). Retrieved Sep. 8, 2010.
- Hoffmeister et al., *J. Biol. Chem.* 280:4329-38 (2005).
- Durre et al. "transcriptional regulation of solventogenesis in *Clostridium acetobutylicum*", *J. Mol Microbiol Biotechnol*, 2002, v. 4, n. 3, p. 295-300.
- Boynton et al "Cloning, sequencing, and expression of clustered genes encoding b-hydroxybutyryl-Coenzyme A (CoA) dehydrogenase, crotonase, and butyryl-coA dehydrogenase from *Clostridium acetobutylicum* ATCC 824" *Journal of Bacteriology*, 1996, v. 178, n. 11, p. 3015-3024.
- Lopez de Felipe, F., et al "Cofactor Engineering: a Novel Approach to Metabolic Engineering in *Lactococcus lactis* by Controlled Expression of NADH Oxidase," *Journal of Bacteriology* 180 (15): 3804-08, American Society for Microbiology, United States (1998).
- Stim-Herndon, K.P., et al., "Characterization of an acetyl-CoA C-acetyltransferase (thiolase) gene from *Clostridium acetobutylicum* ATCC 824," *Gene* 154: 81-85, Elsevier Science B.V., Netherlands (1995).
- "*Escherichia coli* K-12 substr. MG1655 Enzyme: acetyl-CoA acetyltransferase," EcoCyc Accession No. EG11672, accessed on Jan. 24, 2012 at <http://ecocyc.org/ECOLI/NEW-IMAGE?type=GENE&object=EG11672>.
- Casadei, M.A., et al., "Heat resistance of *Bacillus cereus*, *Salmonella typhimurium* and *Lactobacillus delbruekii* in relation to pH and ethanol," *International Journal of Food Microbiology* 63: 125-34, Elsevier Science B.V., Netherlands (2001).
- Berovic, M et al., "influence of Temperature and Carbon Dioxide on Fermentation of Cabernet Sauvignon Must," *Food Technol. Biotechnol.* 41(4): 353-59, Elsevier Science B.V., Netherlands (2003).
- Poulson, C., and Stougaard, P., "Purification and properties of *Saccharomyces cerevisiae* acetolactate synthase from recombinant *Escherichia coli*," *Eur. J. Biochem.* 185: 433-39, FEBS, Elsevier Science B.V., Netherlands (1989).
- Renna: M.C., et al., "Regulation of the *Bacillus subtilis* *alsS*, and *alsR*. Genes Involved in Post-Exponential-Phase Production of Acetoin," *Journal of Bacteriology* 175 (12): 3863-75, American Society for Microbiology, United States (1993).
- Amartey, S.A., et al., "Effects of Temperature and Medium Composition on the Ethanol Tolerance of *Bacillus stearothermophilus* LLD-15," *Biotechnology Letters* 13(9): 627-32, Springer-Verlag GmbH, Germany (1991).
- Herrero, A.A., et al., "Development of Ethanol Tolerance in *Clostridium thermocellum*: Effect of Growth Temperature," *Appl. Environ. Microbiol.*, 40: 571-77, American Society for Microbiology, United States (1980).
- Brown, et al., "The Effects of Temperature on the Ethanol Tolerance of the Yeast *Saccharomyces Ulvarum*," *Biotechnol. Lett.* 4: 269-74, Springer-Verlag, Germany (1982).
- Van Uden, N., et al., "Effects of Ethanol on the Temperature Relations of Viability and Growth in Yeast," *CRC Crit. Rev. Biotechnol.* 1: 263-73, Wiley-Liss, United States (1984).
- Harada, R., et al., "On the Butanol-Rich Production in Acetone-Butanol Fermentation of Molasses (Part 2) Temperature," *Hakko Kyokaishi* 20: 155-56 (1962).
- Jones, et al., "Acetone-Butanol Fermentation Revisited," *Microbiol. Rev.* 50: 484-524, American Society for Microbiology, United States (1986).
- Baer, et al., "Effect of Butanol Challenge and Temperature on Lipid Composition and Membrane Fluidity of Butanol-Tolerant *Clostridium acetobutylicum*," *Appl. Environ. Microbiol.* 53: 2854-61, American Society for Microbiology, United States (1987).
- Inui, M., et al., "Expression of *Clostridium acetobutylicum* butanol synthetic genes in *Escherichia coli*," *Appl. Microbiol. Biotechnol* 77: 305-16, Springer-Verlag, Germany (2008).

* cited by examiner



1

FERMENTATIVE PRODUCTION OF FOUR CARBON ALCOHOLS

CROSS-REFERENCE TO RELATED APPLICATION

This application claims priority under 35 U.S.C. §119 from U.S. Provisional Application Ser. No. 60/721,677, filed Sep. 29, 2005, and from U.S. Provisional Application Ser. No. 60/814,470, filed Jun. 16, 2006.

FIELD OF THE INVENTION

The invention relates to the field of industrial microbiology and the production of alcohols. More specifically, 1-butanol is produced via industrial fermentation of a recombinant microorganism.

BACKGROUND OF THE INVENTION

Butanol is an important industrial chemical, useful as a fuel additive, as a feedstock chemical in the plastics industry, and as a foodgrade extractant in the food and flavor industry. Each year 10 to 12 billion pounds of butanol are produced by petrochemical means and the need for this commodity chemical will likely increase.

Methods for the chemical synthesis of 1-butanol are known, such as the Oxo Process, the Reppe Process, and the hydrogenation of crotonaldehyde (*Ullmann's Encyclopedia of Industrial Chemistry*, 6th edition, 2003, Wiley-VCH Verlag GmbH and Co., Weinheim, Germany, Vol. 5, pp. 716-719). These processes use starting materials derived from petrochemicals and are generally expensive and are not environmentally friendly. The production of 1-butanol from plant-derived raw materials would minimize green house gas emissions and would represent an advance in the art.

Methods for producing 1-butanol by biotransformation of other organic chemicals are also known. For example, Muramoto et al. (JP63017695) describe a method for the production of alcohols, including butanol, from aldehydes using strains of *Pseudomonas*. Additionally, Kuehnle et al. (EP 1149918) describe a process for preparing 1-butanol and 2-butanol by the oxidation of hydrocarbons by various strains of *Rhodococcus ruber*.

Methods of producing butanol by fermentation are also known, where the most popular process produces a mixture of acetone, 1-butanol and ethanol and is referred to as the ABE processes (Blaschek et al., U.S. Pat. No. 6,358,717). Acetone-butanol-ethanol (ABE) fermentation by *Clostridium acetobutylicum* is one of the oldest known industrial fermentations, and the pathways and genes responsible for the production of these solvents have been reported (Girbal et al., *Trends in Biotechnology* 16:11-16 (1998)). The actual fermentation, however, has been quite complicated and difficult to control. ABE fermentation has declined continuously since the 1950s, and almost all butanol is now produced via petrochemical routes, as described above. In a typical ABE fermentation, butyric, propionic, lactic and acetic acids are first produced by *C. acetobutylicum*, the culture pH drops and undergoes a metabolic "butterfly" shift, and 1-butanol, acetone, isopropanol and ethanol are then formed. In conventional ABE fermentations, the 1-butanol yield from glucose is low, typically around 15 percent and rarely exceeding 25 percent. Consequently, the 1-butanol concentration in conventional ABE fermentations is usually lower than 1.3 percent.

Attempts to maximize 1-butanol production from the ABE process by the elimination of all the other solvent by-products

2

have not been totally successful, and thus, the process produces significant amounts of acetone which is not useful as a gasoline additive. A process for the fermentative production of butanol where 1-butanol is the sole product would represent an advance in the art.

There is a need, therefore, for an environmentally responsible, cost-effective process for the production of 1-butanol as a single product. The present invention addresses this need through the discovery of a recombinant microbial production host expressing a 1-butanol biosynthetic pathway.

SUMMARY OF THE INVENTION

The invention provides a recombinant microorganism having an engineered 1-butanol biosynthetic pathway. The engineered microorganism may be used for the commercial production of 1-butanol. Accordingly the invention provides a recombinant microbial host cell comprising at least one DNA molecule encoding a polypeptide that catalyzes a substrate to product conversion selected from the group consisting of:

- a) acetyl-CoA to acetoacetyl-CoA
- b) acetoacetyl-CoA to 3-hydroxybutyryl-CoA
- c) 3-hydroxybutyryl-CoA to crotonyl-CoA
- d) crotonyl-CoA to butyryl-CoA
- e) butyryl-CoA to butyraldehyde and
- f) butyraldehyde to 1-butanol;

wherein the at least one DNA molecule is heterologous to said microbial host cell and wherein said microbial host cell produces 1-butanol.

In another embodiment the invention provides a method for the production of 1-butanol comprising:

- i) providing a recombinant microbial host cell comprising at least one DNA molecule encoding a polypeptide that catalyzes a substrate to product conversion selected from the group consisting of:
 - a) acetyl-CoA to acetoacetyl-CoA
 - b) acetoacetyl-CoA to 3-hydroxybutyryl-CoA
 - c) 3-hydroxybutyryl-CoA to crotonyl-CoA
 - d) crotonyl-CoA to butyryl-CoA
 - e) butyryl-CoA to butyraldehyde and
 - f) butyraldehyde to 1-butanol;

wherein the at least one DNA molecule is heterologous to said microbial host cell; and

- ii) contacting the host cell of (i) with a fermentable carbon substrate under conditions whereby 1-butanol is produced.

BRIEF DESCRIPTION OF THE FIGURES AND SEQUENCE DESCRIPTIONS

The invention can be more fully understood from the following detailed description, FIGURE, and the accompanying sequence descriptions, which form a part of this application.

FIG. 1 shows the 1-butanol biosynthetic pathway. The steps labeled "a", "b", "c", "d", "e", and "f" represent the substrate to product conversions described below.

The following sequences conform with 37 C.F.R. 1.821-1.825 ("Requirements for Patent Applications Containing Nucleotide Sequences and/or Amino Acid Sequence Disclosures—the Sequence Rules") and are consistent with World Intellectual Property Organization (WIPO) Standard ST.25 (1998) and the sequence listing requirements of the EPO and PCT (Rules 5.2 and 49.5(a-bis), and Section 208 and Annex C of the Administrative Instructions). The symbols and format used for nucleotide and amino acid sequence data comply with the rules set forth in 37 C.F.R. §1.822.

A Sequence Listing is provided herewith on Compact Disk. The contents of the Compact Disk containing the Sequence Listing are hereby incorporated by reference in compliance with 37 CFR 1.52(e). The Compact Disks are submitted in triplicate and are identical to one another. The disks are labeled "Copy 1—Sequence Listing", "Copy 2—Sequence Listing", and CRF. The disks contain the following file: CL3241 Conv Seq Listing.ST25 having the following size: 177,000 bytes and which was created Sep. 26, 2006.

TABLE 1

Summary of Gene and Protein SEQ ID Numbers		
Description	SEQ ID NO Nucleic acid	SEQ ID NO Peptide
Acetyl-CoA acetyltransferase thlA from <i>Clostridium acetobutylicum</i> ATCC 824	1	2
Acetyl-CoA acetyltransferase thlB from <i>Clostridium acetobutylicum</i> ATCC 824	3	4
Acetyl-CoA acetyltransferase from <i>Escherichia coli</i>	128	129
Acetyl-CoA acetyltransferase from <i>Bacillus subtilis</i>	130	131
Acetyl-CoA acetyltransferase from <i>Saccharomyces cerevisiae</i>	132	133
3-Hydroxybutyryl-CoA dehydrogenase from <i>Clostridium acetobutylicum</i> ATCC 824	5	6
3-Hydroxybutyryl-CoA dehydrogenase from <i>Bacillus subtilis</i>	134	135
3-Hydroxybutyryl-CoA dehydrogenase from <i>Ralstonia eutropha</i>	136	137
3-Hydroxybutyryl-CoA dehydrogenase from <i>Alcaligenes eutrophus</i>	138	139
Crotonase from <i>Clostridium acetobutylicum</i> ATCC 824	7	8
Crotonase from <i>Escherichia coli</i>	140	141
Crotonase from <i>Bacillus subtilis</i>	142	143
Crotonase from <i>Aeromonas caviae</i>	144	145
Putative trans-enoyl CoA reductase from <i>Clostridium acetobutylicum</i> ATCC 824	9	10
Butyryl-CoA dehydrogenase from <i>Euglena gracilis</i>	146	147
Butyryl-CoA dehydrogenase from <i>Streptomyces collinus</i>	148	149
Butyryl-CoA dehydrogenase from <i>Streptomyces coelicolor</i>	150	151
Butyraldehyde dehydrogenase from <i>Clostridium beijerinckii</i> NRRL B594	11	12
Butyraldehyde dehydrogenase from <i>Clostridium acetobutylicum</i>	152	153
Butanol dehydrogenase bdhB from <i>Clostridium acetobutylicum</i> ATCC 824	13	14
Butanol dehydrogenase bdhA from <i>Clostridium acetobutylicum</i> ATCC 824	15	16
Butanol dehydrogenase from <i>Clostridium acetobutylicum</i>	152	153
Butanol dehydrogenase from <i>Escherichia coli</i>	154	155

SEQ ID NOs:17-44 are the nucleotide sequences of oligo-nucleotide primers used to amplify the genes of the 1-butanol biosynthetic pathway.

SEQ ID NOs:45-72 are the nucleotide sequences of oligo-nucleotide primers used for sequencing.

SEQ ID NOs:73-75 are the nucleotide sequences of oligo-nucleotide primers used to construct the transformation vectors described in Example 9.

SEQ ID NO:76 is the nucleotide sequence of the codon-optimized CAC0462 gene, referred to herein as CaTER.

SEQ ID NO:77 is the nucleotide sequence of the codon-optimized EgTER gene, referred to herein as EgTER(opt).

SEQ ID NO:78 is the nucleotide sequence of the codon-optimized ald gene, referred to herein as ald (opt).

SEQ ID NO:79 is the nucleotide sequence of the plasmid pFP988.

SEQ ID NOs:80-127, 160-185, and 190-207 are the nucleic acid sequences of cloning, sequencing, or PCR screening primers used for the cloning, sequencing, or screening of the genes of the 1-butanol biosynthetic pathway described herein, and are more fully described in Tables 4 and 5.

SEQ ID NO:156 is the nucleotide sequence of the cscBKA gene cluster.

SEQ ID NO:157 is the amino acid sequence of sucrose hydrolase (CscA).

SEQ ID NO:158 is the amino acid sequence of D-fructokinase (CscK).

SEQ ID NO:159 is the amino acid sequence of sucrose permease (CscB).

SEQ ID NO:186 is the nucleotide sequence of the codon optimized tery gene described in Example 17.

SEQ ID NO:187 is the amino acid sequence of the butyl-CoA dehydrogenase (ter) encoded by the codon optimized tery gene (SEQ ID NO: 186).

SEQ ID NO:188 is the nucleotide sequence of the codon optimized aldy gene described in Example 17.

SEQ ID NO:189 is the amino acid sequence of the butyraldehyde dehydrogenase (ald) encoded by the codon optimized aldy gene (SEQ ID NO: 188).

SEQ ID NO:208 is the nucleotide sequence of the template DNA used in Example 14.

DETAILED DESCRIPTION OF THE INVENTION

The present invention relates to methods for the production of 1-butanol using recombinant microorganisms. The present invention meets a number of commercial and industrial needs. Butanol is an important industrial commodity chemical with a variety of applications, where its potential as a fuel or fuel additive is particularly significant. Although only a four-carbon alcohol, butanol has an energy content similar to that of gasoline and can be blended with any fossil fuel. Butanol is favored as a fuel or fuel additive as it yields only CO₂ and little or no SO_x or NO_x when burned in the standard internal combustion engine. Additionally butanol is less corrosive than ethanol, the most preferred fuel additive to date.

In addition to its utility as a biofuel or fuel additive, butanol has the potential of impacting hydrogen distribution problems in the emerging fuel cell industry. Fuel cells today are plagued by safety concerns associated with hydrogen transport and distribution. Butanol can be easily reformed for its hydrogen content and can be distributed through existing gas stations in the purity required for either fuel cells or vehicles.

Finally the present invention produces butanol from plant derived carbon sources, avoiding the negative environmental impact associated with standard petrochemical processes for butanol production.

The following definitions and abbreviations are to be used for the interpretation of the claims and the specification.

5

The term “invention” or “present invention” as used herein is a non-limiting term and is not intended to refer to any single embodiment of the particular invention but encompasses all possible embodiments as described in the specification and the claims.

“ABE” is the abbreviation for the Acetone-Butanol-Ethanol fermentation process.

The term “1-butanol biosynthetic pathway” means the enzyme pathway to produce 1-butanol from acetyl-coenzyme A (acetyl-CoA).

The term “acetyl-CoA acetyltransferase” refers to an enzyme that catalyzes the conversion of two molecules of acetyl-CoA to acetoacetyl-CoA and coenzyme A (CoA). Preferred acetyl-CoA acetyltransferases are acetyl-CoA acetyltransferases with substrate preferences (reaction in the forward direction) for a short chain acyl-CoA and acetyl-CoA and are classified as E.C. 2.3.1.9 [Enzyme Nomenclature 1992, Academic Press, San Diego]; although, enzymes with a broader substrate range (E.C. 2.3.1.16) will be functional as well. Acetyl-CoA acetyltransferases are available from a number of sources, for example, *Escherichia coli* (GenBank Nos: NP_416728 (SEQ ID NO:129), NC_000913 (SEQ ID NO:128); NCBI (National Center for Biotechnology Information) amino acid sequence, NCBI nucleotide sequence), *Clostridium acetobutylicum* (GenBank Nos: NP_349476.1 (SEQ ID NO:2), NC_003030 (SEQ ID NO:1); NP_149242 (SEQ ID NO:4), NC_001988 (SEQ ID NO:3), *Bacillus subtilis* (GenBank Nos: NP_390297 (SEQ ID NO:131), NC_000964 (SEQ ID NO:130)), and *Saccharomyces cerevisiae* (GenBank Nos: NP_015297 (SEQ ID NO:133), NC_001148 (SEQ ID NO:132)).

The term “3-hydroxybutyryl-CoA dehydrogenase” refers to an enzyme that catalyzes the conversion of acetoacetyl-CoA to 3-hydroxybutyryl-CoA. 3-Hydroxybutyryl-CoA dehydrogenases may be reduced nicotinamide adenine dinucleotide (NADH)-dependent, with a substrate preference for (S)-3-hydroxybutyryl-CoA or (R)-3-hydroxybutyryl-CoA and are classified as E.C. 1.1.1.35 and E.C. 1.1.1.30, respectively. Additionally, 3-hydroxybutyryl-CoA dehydrogenases may be reduced nicotinamide adenine dinucleotide phosphate (NADPH)-dependent, with a substrate preference for (S)-3-hydroxybutyryl-CoA or (R)-3-hydroxybutyryl-CoA and are classified as E.C. 1.1.1.157 and E.C. 1.1.1.36, respectively. 3-Hydroxybutyryl-CoA dehydrogenases are available from a number of sources, for example, *C. acetobutylicum* (GenBank Nos: NP_349314 (SEQ ID NO:6), NC_003030 (SEQ ID NO:5)), *B. subtilis* (GenBank Nos: AAB09614 (SEQ ID NO:135), U29084 (SEQ ID NO:134)), *Ralstonia eutropha* (GenBank Nos: YP_294481 (SEQ ID NO:137), NC_007347 (SEQ ID NO:136)), and *Alcaligenes eutrophus* (GenBank Nos: AAA21973 (SEQ ID NO:139), J04987 (SEQ ID NO:138)).

The term “crotonase” refers to an enzyme that catalyzes the conversion of 3-hydroxybutyryl-CoA to crotonyl-CoA and H₂O. Crotonases may have a substrate preference for (S)-3-hydroxybutyryl-CoA or (R)-3-hydroxybutyryl-CoA and are classified as E.C. 4.2.1.17 and E.C. 4.2.1.55, respectively. Crotonases are available from a number of sources, for example, *E. coli* (GenBank Nos: NP_415911 (SEQ ID NO:141), NC_000913 (SEQ ID NO:140)), *C. acetobutylicum* (GenBank Nos: NP_349318 (SEQ ID NO:8), NC_003030 (SEQ ID NO:7)), *B. subtilis* (GenBank Nos: CAB13705 (SEQ ID NO:143), Z99113 (SEQ ID NO:142)), and *Aeromonas caviae* (GenBank Nos: BAA21816 (SEQ ID NO:145), D88825 (SEQ ID NO:144)).

The term “butyryl-CoA dehydrogenase” refers to an enzyme that catalyzes the conversion of crotonyl-CoA to

6

butyryl-CoA. Butyryl-CoA dehydrogenases may be either NADH-dependent or NADPH-dependent and are classified as E.C. 1.3.1.44 and E.C. 1.3.1.38, respectively. Butyryl-CoA dehydrogenases are available from a number of sources, for example, *C. acetobutylicum* (GenBank Nos: NP_347102 (SEQ ID NO:10), NC_003030 (SEQ ID NO:9)), *Euglena gracilis* (GenBank Nos: □5EU90 (SEQ ID NO:147), AY741582 (SEQ ID NO:146)), *Streptomyces collinus* (GenBank Nos: AAA92890 (SEQ ID NO:149), U37135 (SEQ ID NO:148)), and *Streptomyces coelicolor* (GenBank Nos: CAA22721 (SEQ ID NO:151), AL939127 (SEQ ID NO:150)).

The term “butyraldehyde dehydrogenase” refers to an enzyme that catalyzes the conversion of butyryl-CoA to butyraldehyde, using NADH or NADPH as cofactor. Butyraldehyde dehydrogenases with a preference for NADH are known as E.C. 1.2.1.57 and are available from, for example, *Clostridium beijerinckii* (GenBank Nos: AAD31841 (SEQ ID NO:12), AF157306 (SEQ ID NO:11)) and *C. acetobutylicum* (GenBank Nos: NP_149325 (SEQ ID NO:153), NC_001988 (SEQ ID NO:152)).

The term “butanol dehydrogenase” refers to an enzyme that catalyzes the conversion of butyraldehyde to 1-butanol, using either NADH or NADPH as cofactor. Butanol dehydrogenases are available from, for example, *C. acetobutylicum* (GenBank Nos: NP_149325 (SEQ ID NO:153), NC_001988 (SEQ ID NO:152); note: this enzyme possesses both aldehyde and alcohol dehydrogenase activity); NP_349891 (SEQ ID NO:14), NC_003030 (SEQ ID NO:13); and NP_349892 (SEQ ID NO:16), NC_003030 (SEQ ID NO:15)) and *E. coli* (GenBank Nos: NP_417484 (SEQ ID NO:155), NC_000913 (SEQ ID NO:154)).

The term “a facultative anaerobe” refers to a microorganism that can grow in both aerobic and anaerobic environments.

The term “carbon substrate” or “fermentable carbon substrate” refers to a carbon source capable of being metabolized by host organisms of the present invention and particularly carbon sources selected from the group consisting of monosaccharides, oligosaccharides, polysaccharides, and one-carbon substrates or mixtures thereof.

The term “gene” refers to a nucleic acid fragment that is capable of being expressed as a specific protein, optionally including regulatory sequences preceding (5' non-coding sequences) and following (3' non-coding sequences) the coding sequence. “Native gene” refers to a gene as found in nature with its own regulatory sequences. “Chimeric gene” refers to any gene that is not a native gene, comprising regulatory and coding sequences that are not found together in nature. Accordingly, a chimeric gene may comprise regulatory sequences and coding sequences that are derived from different sources, or regulatory sequences and coding sequences derived from the same source, but arranged in a manner different than that found in nature. “Endogenous gene” refers to a native gene in its natural location in the genome of an organism. A “foreign” or “heterologous gene” refers to a gene not normally found in the host organism, but that is introduced into the host organism by gene transfer. Foreign genes can comprise native genes inserted into a non-native organism, or chimeric genes. A “transgene” is a gene that has been introduced into the genome by a transformation procedure.

As used herein the term “coding sequence” refers to a DNA sequence that codes for a specific amino acid sequence. “Suitable regulatory sequences” refer to nucleotide sequences located upstream (5' non-coding sequences), within, or downstream (3' non-coding sequences) of a coding sequence, and

which influence the transcription, RNA processing or stability, or translation of the associated coding sequence. Regulatory sequences may include promoters, translation leader sequences, introns, polyadenylation recognition sequences, RNA processing site, effector binding site and stem-loop structure.

The term "promoter" refers to a DNA sequence capable of controlling the expression of a coding sequence or functional RNA. In general, a coding sequence is located 3' to a promoter sequence. Promoters may be derived in their entirety from a native gene, or be composed of different elements derived from different promoters found in nature, or even comprise synthetic DNA segments. It is understood by those skilled in the art that different promoters may direct the expression of a gene in different tissues or cell types, or at different stages of development, or in response to different environmental or physiological conditions. Promoters which cause a gene to be expressed in most cell types at most times are commonly referred to as "constitutive promoters". It is further recognized that since in most cases the exact boundaries of regulatory sequences have not been completely defined, DNA fragments of different lengths may have identical promoter activity.

The term "operably linked" refers to the association of nucleic acid sequences on a single nucleic acid fragment so that the function of one is affected by the other. For example, a promoter is operably linked with a coding sequence when it is capable of affecting the expression of that coding sequence (i.e., that the coding sequence is under the transcriptional control of the promoter). Coding sequences can be operably linked to regulatory sequences in sense or antisense orientation.

The term "expression", as used herein, refers to the transcription and stable accumulation of sense (mRNA) or antisense RNA derived from the nucleic acid fragment of the invention. Expression may also refer to translation of mRNA into a polypeptide.

As used herein the term "transformation" refers to the transfer of a nucleic acid fragment into a host organism, resulting in genetically stable inheritance. Host organisms containing the transformed nucleic acid fragments are referred to as "transgenic" or "recombinant" or "transformed" organisms.

The terms "plasmid", "vector" and "cassette" refer to an extra chromosomal element often carrying genes which are not part of the central metabolism of the cell, and usually in the form of circular double-stranded DNA fragments. Such elements may be autonomously replicating sequences, genome integrating sequences, phage or nucleotide sequences, linear or circular, of a single- or double-stranded DNA or RNA, derived from any source, in which a number of nucleotide sequences have been joined or recombined into a unique construction which is capable of introducing a promoter fragment and DNA sequence for a selected gene product along with appropriate 3' untranslated sequence into a cell. "Transformation cassette" refers to a specific vector containing a foreign gene and having elements in addition to the foreign gene that facilitates transformation of a particular host cell. "Expression cassette" refers to a specific vector containing a foreign gene and having elements in addition to the foreign gene that allow for enhanced expression of that gene in a foreign host.

As used herein the term "codon degeneracy" refers to the nature in the genetic code permitting variation of the nucleotide sequence without effecting the amino acid sequence of an encoded polypeptide. The skilled artisan is well aware of the "codon-bias" exhibited by a specific host cell in usage of

nucleotide codons to specify a given amino acid. Therefore, when synthesizing a gene for improved expression in a host cell, it is desirable to design the gene such that its frequency of codon usage approaches the frequency of preferred codon usage of the host cell.

The term "codon-optimized" as it refers to genes or coding regions of nucleic acid molecules for transformation of various hosts, refers to the alteration of codons in the gene or coding regions of the nucleic acid molecules to reflect the typical codon usage of the host organism without altering the polypeptide encoded by the DNA.

Standard recombinant DNA and molecular cloning techniques used here are well known in the art and are described by Sambrook, J., Fritsch, E. F. and Maniatis, T., *Molecular Cloning: A Laboratory Manual*, Second Edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1989) (hereinafter "Maniatis"); and by Silhavy, T. J., Bennis, M. L. and Enquist, L. W., *Experiments with Gene Fusions*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1984); and by Ausubel, F. M. et al., *Current Protocols in Molecular Biology*, published by Greene Publishing Assoc. and Wiley-Interscience (1987).

The 1-butanol Biosynthetic Pathway

Carbohydrate utilizing microorganisms employ the Embden-Meyerhof-Parnas (EMP) pathway, the Entner-Doudoroff pathway and the pentose phosphate cycle as the central, metabolic routes to provide energy and cellular precursors for growth and maintenance. These pathways have in common the intermediate glyceraldehyde-3-phosphate and, ultimately, pyruvate is formed directly or in combination with the EMP pathway. Subsequently, pyruvate is transformed to acetyl-coenzyme A (acetyl-CoA) via a variety of means, including reaction with the pyruvate dehydrogenase complex, pyruvate-formate lyase, and pyruvate-ferredoxin oxidoreductase. Acetyl-CoA serves as a key intermediate, for example, in generating fatty acids, amino acids and secondary metabolites. The combined reactions of sugar conversion to acetyl-CoA produce energy (e.g. adenosine-5'-triphosphate, ATP) and reducing equivalents (e.g. reduced nicotinamide adenine dinucleotide, NADH, and reduced nicotinamide adenine dinucleotide phosphate, NADPH). NADH and NADPH must be recycled to their oxidized forms (NAD⁺ and NADP⁺, respectively). In the presence of inorganic electron acceptors (e.g. O₂, NO₃⁻ and SO₄²⁻), the reducing equivalents may be used to augment the energy pool; alternatively, a reduced carbon by-product may be formed. The production of ethanol and 1-butanol resulting from the fermentation of carbohydrate are examples of the latter.

This invention enables the production of 1-butanol from carbohydrate sources with recombinant microorganisms by providing a complete 1-butanol biosynthetic pathway from acetyl-CoA to 1-butanol, as shown in FIG. 1. This biosynthetic pathway, generally lacking in the microbial community due to the absence of genes or the lack of appropriate gene regulation, comprises the following substrate to product conversions:

- a) acetyl-CoA to acetoacetyl-CoA, as catalyzed for example by acetyl-CoA acetyltransferase;
- b) acetoacetyl-CoA to 3-hydroxybutyryl-CoA, as catalyzed for example by 3-hydroxybutyryl-CoA dehydrogenase;
- c) 3-hydroxybutyryl-CoA to crotonyl-CoA, as catalyzed for example by crotonase;
- d) crotonyl-CoA to butyryl-CoA, as catalyzed for example by butyryl-CoA dehydrogenase;
- e) butyryl-CoA to butyraldehyde, as catalyzed for example by butyraldehyde dehydrogenase; and
- f) butyraldehyde to 1-butanol, as catalyzed for example by butanol dehydrogenase.

The pathway requires no ATP and generates NAD⁺ and/or NADP⁺, thus, balances with the central metabolic routes that generate acetyl-CoA. The ability of natural organisms to produce 1-butanol by fermentation is rare and exemplified most prominently by *Clostridium beijerinckii* and *Clostridium acetobutylicum*. The gene organization and gene regulation for *Clostridium acetobutylicum* has been described (L. Girbal and P. Soucaille, *Trends in Biotechnology* 216:11-16 (1998)). However, many of these enzyme activities are associated also with alternate pathways, for example, hydrocarbon utilization, fatty acid oxidation, and polyhydroxyalkanoate metabolism. Thus, in providing a recombinant pathway from acetyl-CoA to 1-butanol, there exist a number of choices to fulfill the individual reaction steps, and the person of skill in the art will be able to utilize publicly available sequences to construct the relevant pathways. A listing of a representative number of genes known in the art and useful in the construction of the 1-butanol biosynthetic pathway are listed below in Table 2.

TABLE 2

Sources of 1-Butanol Pathway Genes	
Gene	GenBank Citation
acetyl-CoA acetyltransferase	NC_000913 <i>Escherichia coli</i> K12, complete genome gi 49175990 ref NC_000913.2 [49175990] NC_001988 <i>Clostridium acetobutylicum</i> ATCC 824 plasmid pSOL1, complete sequence gi 15004705 ref NC_001988.2 [15004705] NC_000964 <i>Bacillus subtilis</i> subsp. <i>subtilis</i> str. 168, complete genome gi 50812173 ref NC_000964.2 [50812173] NC_001148 <i>Saccharomyces cerevisiae</i> chromosome XVI, complete chromosome sequence gi 50593503 ref NC_001148.3 [50593503] CP000017 <i>Streptococcus pyogenes</i> MGAS5005, complete genome gi 71852596 gb CP000017.1 [71852596] NC_005773 <i>Pseudomonas syringae</i> pv. <i>phaseolicola</i> 1448A, complete genome gi 71733195 ref NC_005773.3 [71733195] CR931997 <i>Corynebacterium jeikeium</i> K411 complete genome gi 68262661 emb CR931997.1 [68262661]
3-hydroxybutyryl-CoA dehydrogenase	NC_003030 <i>Clostridium acetobutylicum</i> ATCC 824, complete genome gi 15893298 ref NC_003030.1 [15893298] U29084 <i>Bacillus subtilis</i> (mmgA), (mmgB), (mmgC), and citrate synthase III (mmgD) genes, complete cds, and (mmgE) gene, partial cds gi 881603 gb U29084.1 BSU29084[881603] NC_007347 <i>Ralstonia eutropha</i> JMP134 Raet01_1, whole genome shotgun sequence gi 45517296 ref NZ_AADY01000001.1 [45517296] J04987 <i>A. eutrophus</i> beta-ketothiolase (phbA) and acetoacetyl-CoA reductase (phbB) genes, complete cds gi 141953 gb J04987.1 AFAKTLAACA[141953] NC_004129 <i>Pseudomonas fluorescens</i> Pf-5, complete genome gi 70728250 ref NC_004129.6 [70728250] NC_000913 <i>Escherichia coli</i> K12, complete genome gi 49175990 ref NC_000913.2 [49175990] NC_004557 <i>Clostridium tetani</i> E88, complete genome gi 28209834 ref NC_004557.1 [28209834] NC_006350 <i>Burkholderia pseudomallei</i> K96243 chromosome 1, complete sequence gi 53717639 ref NC_006350.1 [53717639] NC_002947 <i>Pseudomonas putida</i> KT2440, complete genome gi 26986745 ref NC_002947.3 [26986745] NC_000913 <i>Escherichia coli</i> K12, complete genome
crotonase	

TABLE 2-continued

Sources of 1-Butanol Pathway Genes	
Gene	GenBank Citation
	gi 49175990 ref NC_000913.2 [49175990] NC_003030 <i>Clostridium acetobutylicum</i> ATCC 824, complete genome gi 15893298 ref NC_003030.1 [15893298] Z99113 <i>Bacillus subtilis</i> complete genome (section 10 of 21): from 1807106 to 2014934 gi 32468758 emb Z99113.2 BSUB0010[32468758] D88825 <i>Aeromonas caviae</i> phaC gene for PHA synthase, complete cds gi 2335048 dbj D88825.1 [2335048] NC_006274 <i>Bacillus cereus</i> ZK, complete genome gi 52140164 ref NC_006274.1 [52140164] NC_004557 <i>Clostridium tetani</i> E88, complete genome gi 28209834 ref NC_004557.1 [28209834] NC_003030 <i>Clostridium acetobutylicum</i> ATCC 824, complete genome gi 15893298 ref NC_003030.1 [15893298] AY741582 <i>Euglena gracilis</i> trans-2-enoyl-CoA reductase mRNA, complete cds gi 58201539 gb AY741582.1 [58201539] U37135 <i>Streptomyces collinus</i> crotonyl-CoA reductase (ccr) gene, complete cds gi 1046370 gb U37135.1 SCU37135[1046370] AL939127 <i>Streptomyces coelicolor</i> A3(2) complete genome; segment 24/29 gi 24429552 emb AL939127.1 SCO939127[24429552] AP006716 <i>Staphylococcus haemolyticus</i> JCSC1435, complete genome gi 68445725 dbj AP006716.1 [68445725] NC_006274 <i>Bacillus cereus</i> ZK, complete genome gi 52140164 ref NC_006274.1 [52140164] NC_004557 <i>Clostridium tetani</i> E88, complete genome gi 28209834 ref NC_004557.1 [28209834] AF157306 <i>Clostridium beijerinckii</i> strain NRRL B593 hypothetical protein, coenzyme A acylating aldehyde dehydrogenase (ald), acetoacetate:butyrate/acetate coenzyme A transferase (ctfA), acetoacetate:butyrate/acetate coenzyme A transferase (ctfB), and acetoacetate decarboxylase (adc) genes, complete cds gi 47422980 gb AF157306.2 [47422980] NC_001988 <i>Clostridium acetobutylicum</i> ATCC 824 plasmid pSOL1, complete sequence gi 15004705 ref NC_001988.2 [15004705] AY251646 <i>Clostridium saccharoperbutylacetonicum</i> sol operon, complete sequence gi 31075382 gb AY251646.1 [31075382] NC_001988 <i>Clostridium acetobutylicum</i> ATCC 824 plasmid pSOL1, complete sequence gi 15004705 ref NC_001988.2 [15004705] NC_003030 <i>Clostridium acetobutylicum</i> ATCC 824, complete genome gi 15893298 ref NC_003030.1 [15893298] NC_000913 <i>Escherichia coli</i> K12, complete genome gi 49175990 ref NC_000913.2 [49175990] NC_003198 <i>Salmonella enterica</i> subsp. <i>enterica</i> serovar Typhi str. CT18, complete genome gi 16758993 ref NC_003198.1 [16758993] BX571966 <i>Burkholderia pseudomallei</i> strain K96243, chromosome 2, complete sequence gi 52211453 emb BX571966.1 [52211453] Z99120 <i>Bacillus subtilis</i> complete genome (section 17 of 21): from 3213330 to 3414388 gi 32468813 emb Z99120.2 BSUB0017[32468813] NC_003366 <i>Clostridium perfringens</i> str. 13, complete genome gi 18308982 ref NC_003366.1 [18308982] NC_004431 <i>Escherichia coli</i> CFT073, complete genome gi 26245917 ref NC_004431.1 [26245917]
butyryl-CoA dehydrogenase	
butyraldehyde dehydrogenase	
butanol dehydrogenase	

Microbial Hosts for 1-butanol Production

Microbial hosts for 1-butanol production may be selected from bacteria, cyanobacteria, filamentous fungi and yeasts. The microbial host used for 1-butanol production is preferably tolerant to 1-butanol so that the yield is not limited by butanol toxicity. Microbes that are metabolically active at high titer levels of 1-butanol are not well known in the art. Although butanol-tolerant mutants have been isolated from solventogenic *Clostridia*, little information is available concerning the butanol tolerance of other potentially useful bacterial strains. Most of the studies on the comparison of alcohol tolerance in bacteria suggest that butanol is more toxic than ethanol (de Cavalho et al., *Microsc. Res. Tech.* 64:215-22 (2004) and Kabelitz et al., *FEMS Microbiol. Lett.* 220:223-227 (2003)). Tomas et al. (*J. Bacteriol.* 186:2006-2018 (2004)) report that the yield of butanol during fermentation in *Clostridium acetobutylicum* may be limited by butanol toxicity. The primary effect of butanol on *Clostridium acetobutylicum* is disruption of membrane functions (Hermann et al., *Appl. Environ. Microbiol.* 50:1238-1243 (1985)).

The microbial hosts selected for the production of 1-butanol are preferably tolerant to 1-butanol and are able to convert carbohydrates to 1-butanol. The criteria for selection of suitable microbial hosts include the following: intrinsic tolerance to 1-butanol, high rate of glucose utilization, availability of genetic tools for gene manipulation, and the ability to generate stable chromosomal alterations.

Suitable host strains with a tolerance for 1-butanol may be identified by screening based on the intrinsic tolerance of the strain. The intrinsic tolerance of microbes to 1-butanol may be measured by determining the concentration of 1-butanol that is responsible for 50% inhibition of the growth rate (IC₅₀) when grown in a minimal medium. The IC₅₀ values may be determined using methods known in the art. For example, the microbes of interest may be grown in the presence of various amounts of 1-butanol and the growth rate monitored by measuring the optical density at 600 nanometers. The doubling time may be calculated from the logarithmic part of the growth curve and used as a measure of the growth rate. The concentration of 1-butanol that produces 50% inhibition of growth may be determined from a graph of the percent inhibition of growth versus the 1-butanol concentration. Preferably, the host strain should have an IC₅₀ for 1-butanol of greater than about 0.5% weight/volume.

The microbial host for 1-butanol production should also utilize glucose at a high rate. Most microbes are capable of utilizing carbohydrates. However, certain environmental microbes cannot utilize carbohydrates to high efficiency, and therefore would not be suitable hosts.

The ability to genetically modify the host is essential for the production of any recombinant microorganism. The mode of gene transfer technology may be by electroporation, conjugation, transduction or natural transformation. A broad range of host conjugative plasmids and drug resistant markers are available. The cloning vectors are tailored to the host organisms based on the nature of antibiotic resistance markers that can function in that host.

The microbial host also has to be manipulated in order to inactivate competing pathways for carbon flow by deleting various genes. This requires the availability of either transposons to direct inactivation or chromosomal integration vectors. Additionally, the production host should be amenable to chemical mutagenesis so that mutations to improve intrinsic 1-butanol tolerance may be obtained.

Based on the criteria described above, suitable microbial hosts for the production of 1-butanol include, but are not limited to, members of the genera *Clostridium*, *Zymomonas*,

Escherichia, *Salmonella*, *Rhodococcus*, *Pseudomonas*, *Bacillus*, *Lactobacillus*, *Enterococcus*, *Alcaligenes*, *Klebsiella*, *Paenibacillus*, *Arthrobacter*, *Corynebacterium*, *Brevibacterium*, *Pichia*, *Candida*, *Hansenula* and *Saccharomyces*. Preferred hosts include: *Escherichia coli*, *Alcaligenes eutrophus*, *Bacillus licheniformis*, *Paenibacillus macerans*, *Rhodococcus erythropolis*, *Pseudomonas putida*, *Lactobacillus plantarum*, *Enterococcus faecium*, *Enterococcus gallinarum*, *Enterococcus faecalis*, *Bacillus subtilis* and *Saccharomyces cerevisiae*.

Construction of Production Host

Recombinant organisms containing the necessary genes that will encode the enzymatic pathway for the conversion of a fermentable carbon substrate to 1-butanol may be constructed using techniques well known in the art. In the present invention, genes encoding the enzymes of the 1-butanol biosynthetic pathway, i.e., acetyl-CoA acetyltransferase, 3-hydroxybutyryl-CoA dehydrogenase, crotonase, butyryl-CoA dehydrogenase, butyraldehyde dehydrogenase, and butanol dehydrogenase, may be isolated from various sources, as described above.

Methods of obtaining desired genes from a bacterial genome are common and well known in the art of molecular biology. For example, if the sequence of the gene is known, suitable genomic libraries may be created by restriction endonuclease digestion and may be screened with probes complementary to the desired gene sequence. Once the sequence is isolated, the DNA may be amplified using standard primer-directed amplification methods such as polymerase chain reaction (Mullis, U.S. Pat. No. 4,683,202) to obtain amounts of DNA suitable for transformation using appropriate vectors. Tools for codon optimization for expression in a heterologous host are readily available. Some tools for codon optimization are available based on the GC content of the host organism. The GC content of some exemplary microbial hosts is given Table 3.

TABLE 3

GC Content of Microbial Hosts	
Strain	% GC
<i>B. licheniformis</i>	46
<i>B. subtilis</i>	42
<i>C. acetobutylicum</i>	37
<i>E. coli</i>	50
<i>P. putida</i>	61
<i>A. eutrophus</i>	61
<i>Paenibacillus macerans</i>	51
<i>Rhodococcus erythropolis</i>	62
<i>Brevibacillus</i>	50
<i>Paenibacillus polymyxa</i>	50

Once the relevant pathway genes are identified and isolated they may be transformed into suitable expression hosts by means well known in the art. Vectors or cassettes useful for the transformation of a variety of host cells are common and commercially available from companies such as EPICENTRE® (Madison, Wis.), Invitrogen Corp. (Carlsbad, Calif.), Stratagene (La Jolla, Calif.), and New England Biolabs, Inc. (Beverly, Mass.). Typically, the vector or cassette contains sequences directing transcription and translation of the relevant gene, a selectable marker, and sequences allowing autonomous replication or chromosomal integration. Suitable vectors comprise a region 5' of the gene which harbors transcriptional initiation controls and a region 3' of the DNA fragment which controls transcriptional termination. Both control regions may be derived from genes homologous to the

transformed host cell, although it is to be understood that such control regions may also be derived from genes that are not native to the specific species chosen as a production host.

Initiation control regions or promoters, which are useful to drive expression of the relevant pathway coding regions in the desired host cell are numerous and familiar to those skilled in the art. Virtually any promoter capable of driving these genetic elements is suitable for the present invention including, but not limited to, CYC1, HIS3, GAL1, GAL10, ADH1, PGK, PHO5, GAPDH, ADC1, TRP1, URA3, LEU2, ENO, TPI, CUP1, FBA, GPD, and GPM (useful for expression in *Saccharomyces*); AOX1 (useful for expression in *Pichia*); and lac, ara, tet, trp, IP_L , IP_R , T7, tac, and trc (useful for expression in *Escherichia coli*, *Alcaligenes*, and *Pseudomonas*); the amy, apr, npr promoters and various phage promoters useful for expression in *Bacillus subtilis*, *Bacillus licheniformis*, and *Paenibacillus macerans*; nisA (useful for expression Gram-positive bacteria, Eichenbaum et al. *Appl. Environ. Microbiol.* 64(8):2763-2769 (1998)); and the synthetic P11 promoter (useful for expression in *Lactobacillus plantarum*, Rud et al., *Microbiology* 152:1011-1019 (2006)).

Termination control regions may also be derived from various genes native to the preferred hosts. Optionally, a termination site may be unnecessary, however, it is most preferred if included.

Certain vectors are capable of replicating in a broad range of host bacteria and can be transferred by conjugation. The complete and annotated sequence of pRK404 and three related vectors-pRK437, pRK442, and pRK442(H) are available. These derivatives have proven to be valuable tools for genetic manipulation in Gram-negative bacteria (Scott et al., *Plasmid* 50(1):74-79 (2003)). Several plasmid derivatives of broad-host-range Inc P4 plasmid RSF1010 are also available with promoters that can function in a range of Gram-negative bacteria. Plasmid pAYC36 and pAYC37, have active promoters along with multiple cloning sites to allow for the heterologous gene expression in Gram-negative bacteria.

Chromosomal gene replacement tools are also widely available. For example, a thermosensitive variant of the broad-host-range replicon pWV101 has been modified to construct a plasmid pVE6002 which can be used to create gene replacement in a range of Gram-positive bacteria (Maguin et al., *J. Bacteriol.* 174(17):5633-5638 (1992)). Additionally, in vitro transposomes are available to create random mutations in a variety of genomes from commercial sources such as EPICENTRE®.

The expression of the 1-butanol biosynthetic pathway in various preferred microbial hosts is described in more detail below.

Expression of the 1-butanol Biosynthetic Pathway in *E. coli*

Vectors or cassettes useful for the transformation of *E. coli* are common and commercially available from the companies listed above. For example, the genes of the 1-butanol biosynthetic pathway may be isolated from various strains of *Clostridium*, cloned into a modified pUC19 vector and transformed into *E. coli* NM522, as described in Example 11. The expression of the 1-butanol biosynthetic pathway in several other strains of *E. coli* is described in Example 13.

Expression of the 1-butanol Biosynthetic Pathway in *Rhodococcus erythropolis*

A series of *E. coli*-*Rhodococcus* shuttle vectors are available for expression in *R. erythropolis*, including, but not limited to pRhBR17 and pDA71 (Kostichka et al., *Appl. Microbiol. Biotechnol.* 62:61-68 (2003)). Additionally, a series of promoters are available for heterologous gene expression in *R. erythropolis* (see for example Nakashima et al., *Appl.*

Envir. Microbiol. 70:5557-5568 (2004), and Tao et al., *Appl. Microbiol. Biotechnol.* 2005, DOI 10.1007/s00253-005-0064). Targeted gene disruption of chromosomal genes in *R. erythropolis* may be created using the method described by Tao et al., supra, and Brans et al. (*Appl. Envir. Microbiol.* 66: 2029-2036 (2000)).

The heterologous genes required for the production of 1-butanol, as described above, may be cloned initially in pDA71 or pRhBR71 and transformed into *E. coli*. The vectors may then be transformed into *R. erythropolis* by electroporation, as described by Kostichka et al., supra. The recombinants may be grown in synthetic medium containing glucose and the production of 1-butanol can be followed using methods known in the art.

Expression of the 1-butanol Biosynthetic Pathway in *Bacillus subtilis*

Methods for gene expression and creation of mutations in *B. subtilis* are also well known in the art. For example, the genes of the 1-butanol biosynthetic pathway may be isolated from various strains of *Clostridium*, cloned into a modified pUC19 vector and transformed into *Bacillus subtilis* BE1010, as described in Example 12. Additionally, the six genes of the 1-biosynthetic pathway can be split into two operons for expression, as described in Example 14. The first three genes of the pathway (thl, hbd, and crt) were integrated into the chromosome of *Bacillus subtilis* BE1010 (Payne and Jackson, *J. Bacteriol.* 173:2278-2282 (1991)). The last three genes (EgTER, ald, and bdhB) were cloned into expression plasmids and transformed into the *Bacillus* strain carrying the integrated 1-butanol genes.

Expression of the 1-butanol Biosynthetic Pathway in *Bacillus licheniformis*

Most of the plasmids and shuttle vectors that replicate in *B. subtilis* may be used to transform *B. licheniformis* by either protoplast transformation or electroporation. For example, the genes required for the production of 1-butanol may be cloned in plasmids pBE20 or pBE60 derivatives (Nagarajan et al., *Gene* 114:121-126 (1992)). Methods to transform *B. licheniformis* are known in the art (for example see Fleming et al. *Appl. Environ. Microbiol.*, 61(11):3775-3780 (1995)). The plasmids constructed for expression in *B. subtilis* may also be transformed into *B. licheniformis* to produce a recombinant microbial host that produces 1-butanol.

Expression of the 1-butanol Biosynthetic Pathway in *Paenibacillus macerans*

Plasmids may be constructed as described above for expression in *B. subtilis* and used to transform *Paenibacillus macerans* by protoplast transformation to produce a recombinant microbial host that produces 1-butanol.

Expression of the 1-butanol Biosynthetic Pathway in *Alcaligenes (Ralstonia) eutrophus*

Methods for gene expression and creation of mutations in *Ralstonia eutrophus* are known in the art (see for example Taghavi et al., *Appl. Environ. Microbiol.*, 60(10):3585-3591 (1994)). The genes for the 1-butanol biosynthetic pathway may be cloned in any of the broad host range vectors described above, and electroporated to generate recombinants that produce 1-butanol. The polyhydroxy butyrate pathway in *Ralstonia* has been described in detail and a variety of genetic techniques to modify the *Ralstonia eutrophus* genome is known, and those tools can be applied for engineering the 1-butanol biosynthetic pathway.

Expression of the 1-butanol Biosynthetic Pathway in *Pseudomonas putida*

Methods for gene expression in *Pseudomonas putida* are known in the art (see for example Ben-Bassat et al., U.S. Pat. No. 6,586,229, which is incorporated herein by reference).

For example, the butanol pathway genes may be inserted into pPCU18 and this ligated DNA may be electroporated into electrocompetent *Pseudomonas putida* DOT-T1 C5aAR1 cells to generate recombinants that produce 1-butanol.

Expression of the 1-butanol Biosynthetic Pathway in *Saccharomyces cerevisiae*

Methods for gene expression in *Saccharomyces cerevisiae* are known in the art (see for example *Methods in Enzymology*, Volume 194, *Guide to Yeast Genetics and Molecular and Cell Biology* (Part A, 2004, Christine Guthrie and Gerald R. Fink (Eds.), Elsevier Academic Press, San Diego, Calif.). Expression of genes in yeast typically requires a promoter, followed by the gene of interest, and a transcriptional terminator. A number of yeast promoters can be used in constructing expression cassettes for genes encoding the 1-butanol biosynthetic pathway, including, but not limited to constitutive promoters FBA, GPD, and GPM, and the inducible promoters GAL1, GAL10, and CUP1. Suitable transcriptional terminators include, but are not limited to FBAt, GPDt, GPMt, ERG10t, and GAL1t. Suitable promoters, transcriptional terminators, and the genes of the 1-butanol biosynthetic pathway may be cloned into yeast 2 micron (2 μ) plasmids, as described in Example 17.

Expression of the 1-butanol Biosynthetic Pathway in *Lactobacillus plantarum*

The *Lactobacillus* genus belongs to the Lactobacillales family and many plasmids and vectors used in the transformation of *Bacillus subtilis* and *Streptococcus* may be used for *lactobacillus*. Non-limiting examples of suitable vectors include pAM β 11 and derivatives thereof (Renault et al., *Gene* 183:175-182 (1996); and O'Sullivan et al., *Gene* 137:227-231 (1993)); pMBB1 and pHW800, a derivative of pMBB1 (Wyckoff et al. *Appl. Environ. Microbiol.* 62:1481-1486 (1996)); pMG1, a conjugative plasmid (Tanimoto et al., *J. Bacteriol.* 184:5800-5804 (2002)); pNZ9520 (Kleerebezem et al., *Appl. Environ. Microbiol.* 63:45814584 (1997)); pAM401 (Fujimoto et al., *Appl. Environ. Microbiol.* 67:1262-1267 (2001)); and pAT392 (Arthur et al., *Antimicrob. Agents Chemother.* 38:1899-1903 (1994)). Several plasmids from *Lactobacillus plantarum* have also been reported (e.g., van Kranenburg R, Golic N, Bongers R, Leer R J, de Vos W M, Siezen R J, Kleerebezem M. *Appl. Environ. Microbiol.* 2005 March; 71(3): 1223-1230). For example, expression of the 1-butanol biosynthetic pathway in *Lactobacillus plantarum* is described in Example 18.

Expression of the 1-butanol Biosynthetic Pathway in *Enterococcus faecium*, *Enterococcus gallinarum*, and *Enterococcus faecalis*

The *Enterococcus* genus belongs to the Lactobacillales family and many plasmids and vectors used in the transformation of *Lactobacillus*, *Bacillus subtilis*, and *Streptococcus* may be used for *Enterococcus*. Non-limiting examples of suitable vectors include pAM β 1 and derivatives thereof (Renault et al., *Gene* 183:175-182 (1996); and O'Sullivan et al., *Gene* 137:227-231 (1993)); pMBB1 and pHW800, a derivative of pMBB1 (Wyckoff et al. *Appl. Environ. Microbiol.* 62:1481-1486 (1996)); pMG1, a conjugative plasmid (Tanimoto et al., *J. Bacteriol.* 184:5800-5804 (2002)); pNZ9520 (Kleerebezem et al., *Appl. Environ. Microbiol.* 63:45814584 (1997)); pAM401 (Fujimoto et al., *Appl. Environ. Microbiol.* 67:1262-1267 (2001)); and pAT392 (Arthur et al., *Antimicrob. Agents Chemother.* 38:1899-1903 (1994)). Expression vectors for *E. faecalis* using the nisA gene from *Lactococcus* may also be used (Eichenbaum et al., *Appl. Environ. Microbiol.* 64:2763-2769 (1998)). Additionally, vectors for gene replacement in the *E. faecium* chromosome may be used (Nallaapareddy et al., *Appl. Environ. Microbiol.*

72:334-345 (2006)). For example, expression of the 1-butanol biosynthetic pathway in *Enterococcus faecalis* is described in Example 19.

Fermentation Media

Fermentation media in the present invention must contain suitable carbon substrates. Suitable substrates may include but are not limited to monosaccharides such as glucose and fructose, oligosaccharides such as lactose or sucrose, polysaccharides such as starch or cellulose or mixtures thereof and unpurified mixtures from renewable feedstocks such as cheese whey permeate, cornsteep liquor, sugar beet molasses, and barley malt. Additionally the carbon substrate may also be one-carbon substrates such as carbon dioxide, or methanol for which metabolic conversion into key biochemical intermediates has been demonstrated. In addition to one and two carbon substrates methylophilic organisms are also known to utilize a number of other carbon containing compounds such as methylamine, glucosamine and a variety of amino acids for metabolic activity. For example, methylophilic yeast are known to utilize the carbon from methylamine to form trehalose or glycerol (Bellion et al., *Microb. Growth C1 Compd.*, [Int. Symp.], 7th (1993), 415-32. Editor(s): Murrell, J. Collin; Kelly, Don P. Publisher: Intercept, Andover, UK). Similarly, various species of *Candida* will metabolize alanine or oleic acid (Sulter et al., *Arch. Microbiol.* 153:485-489 (1990)). Hence it is contemplated that the source of carbon utilized in the present invention may encompass a wide variety of carbon containing substrates and will only be limited by the choice of organism.

Although it is contemplated that all of the above mentioned carbon substrates and mixtures thereof are suitable in the present invention, preferred carbon substrates are glucose, fructose, and sucrose.

In addition to an appropriate carbon source, fermentation media must contain suitable minerals, salts, cofactors, buffers and other components, known to those skilled in the art, suitable for the growth of the cultures and promotion of the enzymatic pathway necessary for 1-butanol production. Culture Conditions

Typically cells are grown at a temperature in the range of about 25° C. to about 40° C. in an appropriate medium. Suitable growth media in the present invention are common commercially prepared media such as Luria Bertani (LB) broth, Sabouraud Dextrose (SD) broth or Yeast medium (YM) broth. Other defined or synthetic growth media may also be used and the appropriate medium for growth of the particular microorganism will be known by one skilled in the art of microbiology or fermentation science. The use of agents known to modulate catabolite repression directly or indirectly, e.g., cyclic adenosine 2':3'-monophosphate, may also be incorporated into the fermentation medium.

Suitable pH ranges for the fermentation are between pH 5.0 to pH 9.0, where pH 6.0 to pH 8.0 is preferred as the initial condition.

Fermentations may be performed under aerobic or anaerobic conditions, where anaerobic or microaerobic conditions are preferred.

The amount of 1-butanol produced in the fermentation medium can be determined using a number of methods known in the art, for example, high performance liquid chromatography (HPLC) or gas chromatography (GC).

Industrial Batch and Continuous Fermentations

The present process employs a batch method of fermentation. A classical batch fermentation is a closed system where the composition of the medium is set at the beginning of the fermentation and not subject to artificial alterations during the fermentation. Thus, at the beginning of the fermentation the

medium is inoculated with the desired organism or organisms, and fermentation is permitted to occur without adding anything to the system. Typically, however, a "batch" fermentation is batch with respect to the addition of carbon source and attempts are often made at controlling factors such as pH and oxygen concentration. In batch systems the metabolite and biomass compositions of the system change constantly up to the time the fermentation is stopped. Within batch cultures cells moderate through a static lag phase to a high growth log phase and finally to a stationary phase where growth rate is diminished or halted. If untreated, cells in the stationary phase will eventually die. Cells in log phase generally are responsible for the bulk of production of end product or intermediate.

A variation on the standard batch system is the Fed-Batch system. Fed-Batch fermentation processes are also suitable in the present invention and comprise a typical batch system with the exception that the substrate is added in increments as the fermentation progresses. Fed-Batch systems are useful when catabolite repression is apt to inhibit the metabolism of the cells and where it is desirable to have limited amounts of substrate in the media. Measurement of the actual substrate concentration in Fed-Batch systems is difficult and is therefore estimated on the basis of the changes of measurable factors such as pH, dissolved oxygen and the partial pressure of waste gases such as CO₂. Batch and Fed-Batch fermentations are common and well known in the art and examples may be found in Thomas D. Brock in *Biotechnology: A Textbook of Industrial Microbiology*, Second Edition (1989) Sinauer Associates, Inc., Sunderland, Mass., or Deshpande, Mukund V., *Appl. Biochem. Biotechnol.*, 36:227, (1992), herein incorporated by reference.

Although the present invention is performed in batch mode it is contemplated that the method would be adaptable to continuous fermentation methods. Continuous fermentation is an open system where a defined fermentation medium is added continuously to a bioreactor and an equal amount of conditioned media is removed simultaneously for processing. Continuous fermentation generally maintains the cultures at a constant high density where cells are primarily in log phase growth.

Continuous fermentation allows for the modulation of one factor or any number of factors that affect cell growth or end product concentration. For example, one method will maintain a limiting nutrient such as the carbon source or nitrogen level at a fixed rate and allow all other parameters to moderate. In other systems a number of factors affecting growth can be altered continuously while the cell concentration, measured by media turbidity, is kept constant. Continuous systems strive to maintain steady state growth conditions and thus the cell loss due to the medium being drawn off must be balanced against the cell growth rate in the fermentation. Methods of modulating nutrients and growth factors for continuous fermentation processes as well as techniques for maximizing the rate of product formation are well known in the art of industrial microbiology and a variety of methods are detailed by Brock, supra.

It is contemplated that the present invention may be practiced using either batch, fed-batch or continuous processes and that any known mode of fermentation would be suitable. Additionally, it is contemplated that cells may be immobilized on a substrate as whole cell catalysts and subjected to fermentation conditions for 1-butanol production. Methods for 1-butanol Isolation from the Fermentation Medium

The bioproducted 1-butanol may be isolated from the fermentation medium using methods known in the art. For

example, solids may be removed from the fermentation medium by centrifugation, filtration, decantation, or the like. Then, the 1-butanol may be isolated from the fermentation medium, which has been treated to remove solids as described above, using methods such as distillation, liquid-liquid extraction, or membrane-based separation. Because 1-butanol forms a low boiling point, azeotropic mixture with water, distillation can only be used to separate the mixture up to its azeotropic composition. Distillation may be used in combination with another separation method to obtain separation around the azeotrope. Methods that may be used in combination with distillation to isolate and purify 1-butanol include, but are not limited to, decantation, liquid-liquid extraction, adsorption, and membrane-based techniques. Additionally, 1-butanol may be isolated using azeotropic distillation using an entrainer (see for example Doherty and Malone, *Conceptual Design of Distillation Systems*, McGraw Hill, N.Y., 2001).

The 1-butanol-water mixture forms a heterogeneous azeotrope so that distillation may be used in combination with decantation to isolate and purify the 1-butanol. In this method, the 1-butanol containing fermentation broth is distilled to near the azeotropic composition. Then, the azeotropic mixture is condensed, and the 1-butanol is separated from the fermentation medium by decantation. The decanted aqueous phase may be returned to the first distillation column as reflux. The 1-butanol-rich decanted organic phase may be further purified by distillation in a second distillation column.

The 1-butanol may also be isolated from the fermentation medium using liquid-liquid extraction in combination with distillation. In this method, the 1-butanol is extracted from the fermentation broth using liquid-liquid extraction with a suitable solvent. The 1-butanol-containing organic phase is then distilled to separate the 1-butanol from the solvent.

Distillation in combination with adsorption may also be used to isolate 1-butanol from the fermentation medium. In this method, the fermentation broth containing the 1-butanol is distilled to near the azeotropic composition and then the remaining water is removed by use of an adsorbent, such as molecular sieves (Aden et al. *Lignocellulosic Biomass to Ethanol Process Design and Economics Utilizing Co-Current Dilute Acid Prehydrolysis and Enzymatic Hydrolysis for Corn Stover*, Report NREL/TP-510-32438, National Renewable Energy Laboratory, June 2002).

Additionally, distillation in combination with pervaporation may be used to isolate and purify the 1-butanol from the fermentation medium. In this method, the fermentation broth containing the 1-butanol is distilled to near the azeotropic composition, and then the remaining water is removed by pervaporation through a hydrophilic membrane (Guo et al., *J. Membr. Sci.* 245, 199-210 (2004)).

EXAMPLES

The present invention is further defined in the following Examples. It should be understood that these Examples, while indicating preferred embodiments of the invention, are given by way of illustration only. From the above discussion and these Examples, one skilled in the art can ascertain the essential characteristics of this invention, and without departing from the spirit and scope thereof, can make various changes and modifications of the invention to adapt it to various uses and conditions.

General Methods

Standard recombinant DNA and molecular cloning techniques used in the Examples are well known in the art and are described by Sambrook, J., Fritsch, E. F. and Maniatis, T.

19

Molecular Cloning: A Laboratory Manual; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, (1989) (Maniatis) and by T. J. Silhavy, M. L. Bannan, and L. W. Enquist, *Experiments with Gene Fusions*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1984) and by Ausubel, F. M. et al., *Current Protocols in Molecular Biology*, pub. by Greene Publishing Assoc. and Wiley-Interscience (1987).

Materials and methods suitable for the maintenance and growth of bacterial cultures are well known in the art. Techniques suitable for use in the following examples may be found as set out in *Manual of Methods for General Bacteriology* (Phillipp Gerhardt, R. G. E. Murray, Ralph N. Costilow, Eugene W. Nester, Willis A. Wood, Noel R. Krieg and G. Briggs Phillips, eds), American Society for Microbiology, Washington, D.C. (1994)) or by Thomas D. Brock in *Biotechnology: A Textbook of Industrial Microbiology*, Second Edition, Sinauer Associates, Inc., Sunderland, Mass. (1989). All reagents, restriction enzymes and materials used for the growth and maintenance of bacterial cells were obtained from Aldrich Chemicals (Milwaukee, Wis.), BD Diagnostic Systems (Sparks, Md.), Life Technologies (Rockville, Md.), or Sigma Chemical Company (St. Louis, Mo.) unless otherwise specified.

The oligonucleotide primers used for cloning in the following Examples are given in Table 4. The primers used to sequence or screen the cloned genes are given in Table 5. All the oligonucleotide primers were synthesized by Sigma-Genosys (Woodlands, Tex.).

TABLE 4

Oligonucleotide Cloning Primers			
Gene	Primer Name	Sequence	SEQ ID Description
crt	N3	CACCATGGAACATAACAATGTCATCCTTG	17 crt forward
crt	N4	CCTCCTATCTATTTTGTAA GCCTTC	18 crt reverse
hbd	N5	CACCATGAAAAAGGTATGTTATAGGT	19 hbd forward
hbd	N6	CATTTGATAATGGGGATTC TTGT	20 hbd reverse
thlA	N7	CACCATGAAAGAAGTTGTATAGCTAGTGC	21 thlA forward
thlA	N8	CTAGCACTTTTCTAGCAATATTGCTG	22 thlA reverse
bdhA	N9	CACCATGCTAAGTTTGTATTCAATAC	23 bdhA forward
bdhA	N10	TTAATAAGATTTTTTAAATATCTCA	24 bdhA reverse
bdhB	N11	CACCATGGTTGATTTCGAATATTCAATACC	25 bdhB forward
bdhB	N12	TTACACAGATTTTTTGAATATTTGT	26 bdhB reverse
thlB	N15	CACCATGAGAGATGTAGTATAGTAAGTGCTG	27 thlB forward
thlB	N16	CCGCAATTGTATCCATATTGAACC	28 thlB reverse

20

TABLE 4-continued

Oligonucleotide Cloning Primers			
Gene	Primer Name	Sequence	SEQ ID Description
CAC0462	N17	CACCATGATAGTAAAAGCA AAGTTTG	29 CAC0462 forward
CAC0462	N21	GCTTAAAGCTTAAACCGCTTCTGGCG	30 CAC0462 reverse
ald	N27F1	CACCATGAATAAGACACACTAATACC	31 ald forward
ald	N28R1	GCCAGACCATCTTTGAAAA TGCGC	32 ald reverse
thlA	N44	CATGCATGCAAGGAGGTTAGTAGAATGAAAGAAG	33 thlA forward
thlA	N45	GTCCTGCAGGCGCGCCCAATACTTTCTAGCACTTTTC	34 thlA reverse
hbd	N42	CATGTCGACAAAGGAGGTCTGTTTAATGAAAAGGTATG	35 hbd forward
hbd	N43	GTCGCATGCCTTGTAACCTATTTTGAA	36 hbd reverse
CAC0462	N68	CATAGATCTGGATCCAAAGGAGGGTGAGGAAATGATAGTAAAG	37 CAC0462 forward
CAC0462	N69	CATGTCGACGTGCAGCCTTTTAAGGTTCT	38 CAC0462 reverse
crt	N38	CATGAATTCACGCGTAAAGGAGGTATTAGTCATGGAAAC	39 crt forward
crt	N39	GTCGGATCCCTTACCTCCTATCTATTTTG	40 crt reverse
ald	N58	CATGCCCGGGGGTCACCA AAGGAGGAATAGTTCATGATATAA	41 ald forward
ald	N59	CATGGTTAACAAGAAGTTAGCCGGCAAGTACA	42 ald reverse
bdhB	N64	CATGGTTAACAAGGAGGGGTTAAATGGTTGATTTCGAAT	43 bdhB forward
bdhB	N65	CATGGCATGCGTTTAAACGTAGGTTTACACAGATTTT	44 bdhB reverse
—	BenF	ACTTTCTTTTCGCCTGTTTCAC	73 —
—	BenMAR	CATGAAGCTTGGCGCGCCGGGACGCTTTTGAATAATGAAAAC	74 —
—	BenBPR	CATGAAGCTTGTTTAACTCGGTGACCTTGAAAATAATGAAACTTATATTG	75 —
EgTER (opt)	N85	CATAGATCTGGATCCAAAGGAGGGTGAGGAAATGGCGATGTTTACG	80 Egter forward
EgTER (opt)	N86	GTCGACTTACTGCTGGGCGG	81 Egter reverse

21

TABLE 4-continued

Oligonucleotide Cloning Primers			
Gene	Primer Name	Sequence	SEQ ID Description
Ptrec-ald(opt)	T-Ptrec(BspEI)	TTCCGTACTTCCGGACGAC TGCACGGTGCACCAATGC TTCTG	87Ptrec forward
Ptrec-ald(opt)	B-ald(opt)(Scal)	CGGATCTTAAGTACTTTAA CCGCCAGCACACAGCGGC GCTGG	88ald reverse
ald	AF BamHI	CATTGGATCCATGAATAAA GACACACTAATACCTACAA C	93ald forward
ald	AR Aat2	CATGACGTCAGTAGTGTTA ACAAGAAGTTAGCCGGCA AG	94ald reverse
EgTER	Forward 1 (E)	CATGTTAAACAAAGGAGGAA AGATCTATGGCGATGTTTA CGACCACCGCAA	95EgTER SOE forward
EgTER	Bottom Reverse 1 (E)	CCCCTCCTTTGGCGCGCC TTACTGCTGGCGGCGCT CGGCAGA	96EgTER SOE reverse
bdh	Top Forward 2 (B)	GCCCAGCAGTAAGCGCGC CCAAAGGAGGGGTTAAAT GGTTGATTTTGAAT	97bdh SOE forward
bdh	Reverse 2 (B)	GTCGACGTCATACTAGTTT ACACAGATTTTTTGAATATT TGT	98bdh SOE reverse
—	Pamy/la cO F	CATTGTACAGAATTCGAGC TCTCGAGGCCCGCACAT ACGAAAAGAC	99Pamy forward
—	Pamy/la cO R	CATTGTACAGTTTAAACAT AGGTCACCCCTCATTTTCGT AGGAATTGTTATCC	100Pamy reverse
—	Spac F	CATCTCGAGTAATCTACA CAGCCCAGTCC	101Pspac forward
—	Spac R	CATGTTTAAACGGTGACCC AAGCTGGGGATCCGCGG	102Pspac reverse
thl	Top TF	CATTGGTCACCATTTCCCGG GCATGCAAAGGAGGTTAG TAGAATG	103thl SOE Forward
thl	Bot TR	CCTTTACGCGACCGGTACT AGTCAAGTCGACAGGGCG CGCCCAATACTTTC	104thl SOE reverse
crt	Top CF	CGCGCCCTGTCGACTTGA CTAGTACCGGTGCGGTAAA GGAGGTATTAGTCATGGAA C	105crt SOE forward
crt	Bot CR	CATCGTTTAAACTTGGATC CAGATCCCTTACCTCCTAT	106crt SOE reverse
ERG10-ERG10t	OT731	AAAGCTGGAGCTCCACCG CGGTGGCGGCGCTCTAG AAGTTTTCAAAGCAGAGTT TCGTTTGAATATTTTACCA	164ERG10-ERG10t forward
ERG10-ERG10t	OT732	TTCAATATGCATGCCTCAG AACGTTTACATTGTATCGA CTGCCAGAACCC	165ERG10-ERG10t reverse

22

TABLE 4-continued

Oligonucleotide Cloning Primers			
Gene	Primer Name	Sequence	SEQ ID Description
GAL1-GAL10	OT733	GCAGTCGATACAATGTAAA CGTTCTGAGGCATGCATAT TGAATTTTCAAAAATTCTTA CTTTTTTTTGGATGGACG CA	166GAL1-GAL10 forward
GAL1-GAL10	OT734	ACCTGCACCTATAACACAT ACCTTTTCCATGGTAGTTT TTTCTCCTTGACGTTAAAG TATAGAGGTATATTA	167GAL1-GAL10 reverse
hbd	OT735	AAAACTACCATGGAAAAG GTATGTGTTATAGTGCAG GTACTATGGGTTCCAGGAAT TGC	168hbd forward
hbd	OT736	GTAAAAAAGAAGGCCCGT ATAGGCCTTATTTTGAATA ATCGTAGAAACCTTTTCCT GATTTTCTTCCAAG	169hbd reverse
GAL1t	OT737	ACGATTATTCAAAATAAGG CCTATACGGCTTCTTTTT TTTACTTTGTTCCAGAACAA CTTCTCATTTTTTCTACTC ATAA	170GAL1t forward
GAL1t	OT738	GAATTGGGTACCGGGCCC CCCCTCGAGGTGCGACCGA TGCTCATAAACTTCGGTA GTTATATTACTCTGAGAT	171GAL1t reverse
thlA	OT797	AAAGTAAGAATTTTTGAAA ATTCAATATGCATGCAAGA AGTTGTAATAGCTAGTGCA GTAAGAAC	172thlA forward
thlA	OT798	GAAAAAGATCATGAGAAAA TCGCAGAACCTAAGGCGC GCCTCAGCACTTTTCTAGC AATATTGCTGTTCTCTTG	173thlA reverse
CUP1	OT806	CTCGAAAAATAGGGCGCGC CCCCATTACCGACATTTGG GCGC	174CUP1 forward
CUP1	OT807	ACTGCACTAGCTATTACAA CTTCTTGCATGCGTGATGA TTGATTGATTGATTGA	175CUP1 reverse
GPD promoter	OT808	TCGGTAATGGGGCGCGC CCTATTTTCGAGGACCTTG TCACCTTGA	176GPD promoter forward
GPD promoter	OT809	TTTCGAATAAACACACATA AACAAACACCCCATGGAAA AGGTATGTGTTATAGGTGC AGG	177GPD promoter reverse
FBA1 promoter	OT799	TACCGGGCCCCCCTCGA GGTCGACGGCGCGCCACT GGTAGAGAGCGACTTTGTA TGCCCCA	178FBA1 promoter forward
FBA1 promoter	OT761	CTTGGCCTTCACTAGCATG CTGAATATGTATTACTTGG TTATGGTTATATATGACAAA AG	179FBA1 promoter reverse
GPM1 promoter	OT803	CCCTCACTAAAGGGAACAA AAGCTGGAGCTCGATATC GGCGCGCCACATGCAGT GATGCACGCGCA	180GPM1 promoter forward

23

TABLE 4-continued

Oligonucleotide Cloning Primers			
Gene	Primer Name	Sequence	SEQ ID Description
GPM1 promoter	OT804	AAGGATGACATTGTTTAGT TCCATGGTTGTAATATGTG TGTTTGTGTTGG	181 GPM1 promoter reverse
crt	OT785	CACACATATTACAACCATG GAACATAAACAATGTCATCC TTGAAAAGGAAGG	182 Crt forward
crt	OT786	ATCATTTCATTGGCCATTCA GGCCTTATCTATTTTGGAA GCCTTCAATTTTCTTTTCT CTATG	183 Crt reverse
GPM1 terminator	OT787	CAAAATAGATAAGGCCTG AATGGCCAATGAATGATTT GATGATTCTTTTCCCTC CATTTTTC	184 GPM1t terminator forward
GPM1t terminator	OT805	GAATTGGGTACCGGGCCC CCCCTCGAGGTCGACTTAT AGTATTATATTTCTGATTT GGTTATAGCAAGCAGCGTT T	185 GPM1t terminator reverse
GPD promoter	OT800	GGGAACAAAAGCTGGAGC TCCACCGCGGTGGGGCGC GCCCTATTTTCGAGGACCT TGTACCTTGAGCC	190 GPD promoter forward
GPD promoter	OT758	TTAAGGTATCTTTATCCAT GGTGTGTTGTTTATGTGTG TTATTCGAAACT	191 GPD promoter reverse
GPD terminator	OT754	TTGGGTACCGGGCCCCC CTCGAGGTCGACTGGCCA TTAATCTTTCCCATAT	192 GPD terminator forward
GPD terminator	OT755	TGTGTCCTAGCAGGTTAGG GCCTGCAGGGCCGTGAAT TTACTTTAAATCTTG	193 GPD terminator reverse
FBA1 promoter	OT760	CGAAAAATAGGCGCGCCA CTGGTAGAGAGCGACTTT GTATGCCCAATTG	194 FBA1 promoter forward
FBA1 promoter	OT792	CCCTTGACGAACCTTGGCCT TCACTAGCATGCTGAATAT GTATTACTTGGTTATGGTT ATATATGACAAAAG	195 FBA1 promoter reverse
FBA1 terminator	OT791	CCCTTGACGAACCTTGGCCT TCACTAGCATGCTGAATAT GTATTACTTGGTTATGGTT ATATATGACAAAAG	196 FBA1 terminator forward
FBA1 terminator	OT765	GGAACAAAAGCTGGAGCT CCACCGCGGTGGTTTAAAC GTATAGACTTCTAATATATT TCTCCATACTTGGTATT	197 FBA1 terminator reverse
ldhL	LDH EcoRV F	GACGTCATGACCACCCGC CGATCCCTTTT	198 ldhL forward
ldhL	LDH AatIIR	GATATCCAACACCGAGC CGACGTATTAC	199 ldhL reverse
Cm	Cm F	ATTTAAATCTCGAGTAGAG GATCCCAACAACGAAAAT TGGATAAAG	200 Cm forward

24

TABLE 4-continued

Oligonucleotide Cloning Primers			
Gene	Primer Name	Sequence	SEQ ID Description
Cm	Cm R	ACGCGTTATTATAAAAGCC AGTCATTAGG	201 Cm reverse
P11	P11 F	TCGAGAGCGCTATAGTTGT TGACAGAATGGACATACTA TGATATATTGTTGCTATAG CGCCC	202 P11 promoter forward
P11	P11 R	GGGCGCTATAGCAACAATA TATCATAGTATGTCATTCT GTCAACAACATATAGCGCTC	203 P11 promoter reverse
PldhL	PldhL F	GAGCTCGTCGACAAACCA ACATTATGACGTGTCTGGG C	204 ldhL promoter forward
PldhL	PldhL R	GGATCCTACCATGTTTGTG CAAAATAAGTG	205 ldhL promoter reverse
PnisA	F-PnisA (EcoRV)	TTCACTGATATCGACATAC TTGAATGACCTAGTC	206 PnisA forward
PnisA	R-PnisA (Pmel BamHI)	TTGATTAGTTTAAACTGTA GGATCCCTTTGAGTGCCTCC TTATAATTTA	207 PnisA reverse

TABLE 5

Sequencing and PCR Screening Primers			
Name	Sequence	Gene-specific	SEQ ID NO:
M13 Forward	GTAAAACGACGGCCAGT	TOPO vector	45
M13 Reverse	AACAGCTATGACCATG	TOPO vector	46
N7SeqF1	GCAGGAGATGCTGACGTAATAA	thlA	47
N7SeqR1	CCAACCTGCTTTTCAATAGCTGC	thlA	48
N15SeqF1	CAGAGATGGGGTCAAGAATG	thlB	49
N16SeqR1	GTGGTTTTATTCCGAGAGCG	thlB	50
N5SeqF2	GGTCTATACTTAGAATCTCC	hbd	51
N6SeqR2	CGGAACAGTTGACCTTAATATGGC	hbd	52
N22SeqF1	GCCTCATCTGGGTTTGGTCTTG	CAC0426	53
N22SeqF2	CGCCTAGGAGAAAGGACTATAAAA CTGG	CAC0426	54
N22SeqF3	CAGAGTTATAGGTGGTAGAGCC	CAC0426	55
N23SeqR1	CCATCCCGCTGTTCTATTCTTCT	CAC0426	56
N23SeqR2	CCAATCCTCTCCACCATACC	CAC0426	57
N23SeqR3	CGTCCATCCTTAATCTTCCC	CAC0426	58
N31SeqF2	CCAATATGGAATCCCTAGATGC	ald	59
N31SeqF3	GCATAGTCTGCGAAGTAAATGC	ald	60

25

TABLE 5-continued

Sequencing and PCR Screening Primers			
Name	Sequence	Gene-specific	SEQ ID NO:
N31SeqF4	GGATCTACTGGTGAAGGCATAACC	ald	61
N32SeqR1	GTTAGCCGGCAAGTACACATC	ald	72
N32SeqR2	GGCATCATGAGTTCTGTCATGAC	ald	62
N32SeqR3	GCCTTCAATGATACTCTTACCAGCCald		63
N32SeqR4	GCATTTCCAGCAGCTATCATGC	ald	64
N32SeqR5	CCTTCCCATATGTGTTTCTTCC	ald	65
N11SeqF1	GTTGAAGTAGTACTAGCTATAG	bdhB	66
N11SeqF2	GACATAACACACGGCGTAGGGC	bdhB	67
N12SeqR1	TAAGTGTACACTCCAATTAGTG	bdhB	68
N12SeqR2	GCCATCTAACACAATATCCCATGG	bdhB	69
N9SeqF1	GCGATACATGGGACATGGTTAAAG	bdhA	70
N10SeqR1	TGCACTTAACCTCGTGTTCATA	bdhA	71
T7Primer	TAATACGACTCACTATAGGG	pET23 vector	82
Trc99aF	TTGACAATTAATCATCCGGC	p Trc99a vector	83
N5SeqF4	GGTCAACTGTTCCGGAAATTC	hbd	84
T-ald (BamHI)	TGATCTGGATCCAAGAAGGAGCCCC TTACCATGAATAAGACACAC	ald	85
B-ald (EgTER)	CATCGCCATTTCTCTCACCCTCCTTTald TTAGCCGGCAAGTACACATCTTCTT TGTC		86
N3SeqF1	CCATCATACCATACTGACCC	crt	107
N3SeqF2	GCTACTGGAGCATGTCTCAC	crt	108
N3SeqF3	CCATTAACAGCTGCTATTACAGGC	crt	109
N4SeqR3	GGTCTCGGAATAACACCTGG	crt	110
N5SeqF3	CAAGCTTCATAACAGGAGCTGG	hbd	111
N7SeqR2	ATCCCAACATCCGTCACTGATC	thlA	112
N31SeqF1	CTGAGATAAGAAAGCCGCA	ald	113
N62SeqF2	CAACCTTGGGCGTGTCTCTG	EgTER	114
N62SeqF3	GTGGCGAAGATTGGGAAGT	EgTER	115
N62SeqF4	GGGAAATGGCAGAAGATGTTGAGC	EgTER	116
N63SeqR1	CGGTCTGATAAAGTCAAAATCGC	EgTER	117
N63SeqR2	CACCAGCGCTTTGGCAACAAC	EgTER	118
N63SeqR3	GAACGTGCATACAGACCTGCTTC	EgTER	119
N63SeqR4	CGGCTGAATAACTTTTGCGG	EgTER	120
Pamy SeqF2	GCCTTTGATGACTGATGATTGGC	pFP988 vector	121
Pamy SeqF	TCTCCGGTAAACATTACGGCAAAC	pFP988 vector	122

26

TABLE 5-continued

Sequencing and PCR Screening Primers			
Name	Sequence	Gene-specific	SEQ ID NO:
Pamy SeqR	CGGTCAGATGCAATTCGACATGTG	pFP988 vector	123
SpacF Seq	GAAGTGGTCAAGACCTCACT	Pspac promoter	124
sacB Up	CGGGTTTGTACTGATAAAGCAGG	sacB	125
sacB Dn	CGGTTAGCCATTTGCCTGCTTTTA	sacB	126
HT R	ACAAAGATCTCCATGGACGCGT	pHT01 vector	127
Scr1	CCTTTCTTTGTGAATCGG	csc	160
Scr2	AGAAACAGGGTGTGATCC	csc	161
Scr3	AGTGATCATCACCTGTTGCC	csc	162
Scr4	AGCACGGCGAGAGTCGACGG	csc	163

25 Methods for Determining 1-butanol Concentration in Culture Media

The concentration of 1-butanol in the culture media can be determined by a number of methods known in the art. For example, a specific high performance liquid chromatography (HPLC) method utilized a Shodex SH-1011 column with a Shodex SH-G guard column, both purchased from Waters Corporation (Milford, Mass.), with refractive index (RI) detection. Chromatographic separation was achieved using 0.01 M H₂SO₄ as the mobile phase with a flow rate of 0.5 mL/min and a column temperature of 50° C. 1-butanol had a retention time of 52.8 min under the conditions used. Alternatively, gas chromatography (GC) methods are available. For example, a specific GC method utilized an HP-INNOWax column (30 m×0.53 mm id, 1 μm film thickness, Agilent Technologies, Wilmington, Del.), with a flame ionization detector (FID). The carrier gas was helium at a flow rate of 4.5 mL/min, measured at 150° C. with constant head pressure; injector split was 1:25 at 200° C.; oven temperature was 45° C. for 1 min, 45 to 220° C. at 10° C./min, and 220° C. for 5 min; and FID detection was employed at 240° C. with 26 mL/min helium makeup gas. The retention time of 1-butanol was 5.4 min. A similar GC method using a Varian CP-WAX 58(FFAP) CB column (25 m×0.25 mm id×0.2 μm film thickness, Varian, Inc., Palo Alto, Calif.) was also used.

The meaning of abbreviations is as follows: “s” means second(s), “min” means minute(s), “h” means hour(s), “psi” means pounds per square inch, “nm” means nanometers, “d” means day(s), “μL” means microliter(s), “mL” means milliliter(s), “L” means liter(s), “mm” means millimeter(s), “nm” means nanometers, “mM” means millimolar, “M” means molar, “mmol” means millimole(s), “μmole” means micromole(s), “g” means gram(s), “μg” means microgram(s) and “ng” means nanogram(s), “PCR” means polymerase chain reaction, “OD” means optical density, “OD₆₀₀” means the optical density measured at a wavelength of 600 nm, OD₅₅₀” means the optical density measured at a wavelength of 550 nm, “kDa” means kilodaltons, “g” means the gravitation constant, “rpm” means revolutions per minute, “bp” means base pair(s), “kbp” means kilobase pair(s), “% w/v” means weight/volume percent, % v/v” means volume/volume percent,

“HPLC” means high performance liquid chromatography, and “GC” means gas chromatography.

Example 1

Cloning and Expression of Acetyl-CoA Acetyltransferase

The purpose of this Example was to express the enzyme acetyl-CoA acetyltransferase, also referred to herein as acetoacetyl-CoA thiolase, in *E. coli*. The acetoacetyl-CoA thiolase gene *thlA* was cloned from *C. acetobutylicum* (ATCC 824) and expressed in *E. coli*. The *thlA* gene was amplified from *C. acetobutylicum* (ATCC 824) genomic DNA using PCR, resulting in a 1.2 kbp product.

The genomic DNA from *Clostridium acetobutylicum* (ATCC 824) was either purchased from the American Type Culture Collection (ATCC, Manassas, Va.) or was isolated from *Clostridium acetobutylicum* (ATCC 824) cultures, as described below.

Genomic DNA from *Clostridium acetobutylicum* (ATCC 824) was prepared from anaerobically grown cultures. The *Clostridium* strain was grown in 10 mL of Clostridial growth medium (Lopez-Contreras et al., *Appl. Env. Microbiol.* 69(2), 869-877 (2003)) in stoppered and crimped 100 mL Bellco serum bottles (Bellco Glass Inc., Vineland, N.J.) in an anaerobic chamber at 30° C. The inoculum was a single colony from a 2xYTG plate (Kishii, et al., *Antimicrobial Agents & Chemotherapy*, 47(1), 77-81 (2003)) grown in a 2.5 L MGC AnaeroPak™ (Mitsubishi Gas Chemical America Inc, New York, N.Y.) at 37° C.

Genomic DNA was prepared using the Gentra Puregene® kit (Gentra Systems, Inc., Minneapolis, Minn.; catalog no. D-6000A) with modifications to the manufacturer's instruction (Wong et al., *Current Microbiology*, 32, 349-356 (1996)). The *thlA* gene was amplified from *Clostridium acetobutylicum* (ATCC 824) genomic DNA by PCR using primers N7 and N8 (see Table 4), given as SEQ ID NOs:21 and 22, respectively. Other PCR amplification reagents were supplied in manufacturers' kits for example, Kod HiFi DNA Polymerase (Novagen Inc., Madison, Wis.; catalog no. 71805-3) and used according to the manufacturer's protocol. Amplification was carried out in a DNA Thermocycler GeneAmp 9700 (PE Applied Biosystems, Foster city, CA).

For expression studies the Gateway cloning technology (Invitrogen Corp., Carlsbad, Calif.) was used. The entry vector pENTR/SD/D-TOPO allowed directional cloning and provided a Shine-Dalgarno sequence for the gene of interest. The destination vector pDEST14 used a T7 promoter for expression of the gene with no tag. The forward primer incorporated four bases (CACC) immediately adjacent to the translational start codon to allow directional cloning into pENTR/SD/D-TOPO (Invitrogen) to generate the plasmid pENTRSDD-TOPOthIA. The pENTR construct was transformed into *E. coli* Top10 (Invitrogen) cells and plated according to manufacturer's recommendations. Transformants were grown overnight and plasmid DNA was prepared using the QIAprep Spin Miniprep kit (Qiagen, Valencia, Calif.; catalog no. 27106) according to manufacturer's recommendations. Clones were submitted for sequencing with M13 Forward and Reverse primers (see Table 5), given as SEQ ID NOs:45 and 46, respectively, to confirm that the genes inserted in the correct orientation and to confirm the sequence. Additional sequencing primers, N7SeqF1 and N7SeqR1 (see Table 5), given as SEQ ID NOs:47 and 48, respectively, were needed to completely sequence the PCR product. The nucleotide sequence of the open reading frame

(ORF) for this gene and the predicted amino acid sequence of the enzyme are given as SEQ ID NO:1 and SEQ ID NO:2, respectively.

To create an expression clone, the *thlA* gene was transferred to the pDEST 14 vector by recombination to generate pDEST14thIA. The pDEST14thIA vector was transformed into BL21-AI cells. Transformants were inoculated into LB medium supplemented with 50 µg/mL of ampicillin and grown overnight. An aliquot of the overnight culture was used to inoculate 50 mL of LB supplemented with 50 µg/mL of ampicillin. The culture was incubated at 37° C. with shaking until the OD₆₀₀ reached 0.6-0.8. The culture was split into two 25-mL cultures and arabinose was added to one of the flasks to a final concentration of 0.2% by weight. The negative control flask was not induced with arabinose. The flasks were incubated for 4 h at 37° C. with shaking. Cells were harvested by centrifugation and the cell pellets were resuspended in 50 mM MOPS, pH 7.0 buffer. The cells were disrupted either by sonication or by passage through a French Pressure Cell. The whole cell lysate was centrifuged yielding the supernatant or cell free extract and the pellet or the insoluble fraction. An aliquot of each fraction (whole cell lysate, cell free extract and insoluble fraction) was resuspended in SDS (MES) loading buffer (Invitrogen), heated to 85° C. for 10 min and subjected to SDS-PAGE analysis (NUPAGE 4-12% Bis-Tris Gel, catalog no. NP0322Box, Invitrogen). A protein of the expected molecular weight of about 41 kDa, as deduced from the nucleic acid sequence, was present in the induced culture but not in the uninduced control.

Acetoacetyl-CoA thiolase activity in the cell free extracts was measured as degradation of a Mg²⁺-acetoacetyl-CoA complex by monitoring the decrease in absorbance at 303 nm. Standard assay conditions were 100 mM Tris-HCl pH 8.0, 1 mM DTT (dithiothreitol) and 10 mM MgCl₂. The cocktail was equilibrated for 5 min at 37° C.; then the cell-free extract was added. The reaction was initiated with the addition of 0.05 mM acetoacetyl-CoA plus 0.2 mM CoA. Protein concentration was measured by either the Bradford method or by the Bicinchoninic Kit (Sigma, catalog no. BCA-1). Bovine serum albumin (Bio-Rad, Hercules, Calif.) was used as the standard in both cases. In one typical assay, the specific activity of the ThIA protein in the induced culture was determined to be 16.0 µmol mg⁻¹ min⁻¹ compared to 0.27 µmol mg⁻¹ min⁻¹ in the uninduced culture.

Example 2

Cloning and Expression of Acetyl-CoA Acetyltransferase

The purpose of this Example was to express the enzyme acetyl-CoA acetyltransferase, also referred to herein as acetoacetyl-CoA thiolase, in *E. coli*. The acetoacetyl-CoA thiolase gene *thlB* was cloned from *C. acetobutylicum* (ATCC 824) and expressed in *E. coli*. The *thlB* gene was amplified from *C. acetobutylicum* (ATCC 824) genomic DNA using PCR.

The *thlB* gene was cloned and expressed in the same manner as the *thlA* gene described in Example 1. The *C. acetobutylicum* (ATCC 824) genomic DNA was amplified by PCR using primers N15 and N16 (see Table 4), given as SEQ ID NOs:27 and 28, respectively, creating a 1.2 kbp product. The forward primer incorporated four bases (CCAC) immediately adjacent to the translational start codon to allow directional cloning into pENTR/SD/D-TOPO (Invitrogen) to generate the plasmid pENTRSDD-TOPOthIB. Clones were submitted for sequencing with M13 Forward and Reverse primers,

29

given as SEQ ID NOs:45 and 46 respectively, to confirm that the genes inserted in the correct orientation and to confirm the sequence. Additional sequencing primers, N15SeqF1 and N16SeqR1 (see Table 5), given as SEQ ID NOs:49 and 50 respectively, were needed to completely sequence the PCR product. The nucleotide sequence of the open reading frame (ORF) for this gene and the predicted amino acid sequence of the enzyme are given as SEQ ID NO:3 and SEQ ID NO:4, respectively.

To create an expression clone, the thlB gene was transferred to the pDEST 14 (Invitrogen) vector by recombination to generate pDEST14thlB. The pDEST14thlB vector was transformed into BL21-AI cells and expression from the T7 promoter was induced by addition of arabinose. A protein of the expected molecular weight of about 42 kDa, as deduced from the nucleic acid sequence, was present in the induced culture, but not in the uninduced control. Enzyme assays were performed as described in Example 1. In one typical assay, the specific activity of the ThlB protein in the induced culture was determined to be 14.9 $\mu\text{mol mg}^{-1} \text{ min}^{-1}$ compared to 0.28/ $\mu\text{mol mg}^{-1} \text{ min}^{-1}$ in the uninduced culture.

Example 3

Cloning and Expression of 3-Hydroxybutyryl-CoA Dehydrogenase

The purpose of this Example was to clone the hbd gene from *C. acetobutylicum* (ATCC 824) and express it in *E. coli*. The hbd gene was amplified from *C. acetobutylicum* (ATCC 824) genomic DNA using PCR.

The hbd gene was cloned and expressed using the method described in Example 1. The hbd gene was amplified from *C. acetobutylicum* (ATCC 824) genomic DNA by PCR using primers N5 and N6 (see Table 4) given as SEQ ID NOs:19 and 20 respectively, creating a 881 bp product. The forward primer incorporated four bases (CACC) immediately adjacent to the translational start codon to allow directional cloning into pENTR/SD/D-TOPO (Invitrogen) to generate the plasmid pENTRSDD-TOPOhbd. Clones were submitted for sequencing with M13 Forward and Reverse primers, given as SEQ ID NOs:45 and 46 respectively, to confirm that the genes inserted in the correct orientation and to confirm the sequence. Additional sequencing primers, N5SeqF2 and N6SeqR2 (see Table 5), given as SEQ ID NOs:51 and 52 respectively, were needed to completely sequence the PCR product. The nucleotide sequence of the open reading frame (ORF) for this gene and the predicted amino acid sequence of the enzyme are given as SEQ ID NO:5 and SEQ ID NO:6, respectively.

To create an expression clone, the hbd gene was transferred to the pDEST 14 (Invitrogen) vector by recombination to generate pDEST14hbd. The pDEST14hbd vector was transformed into BL21-AI cells and expression from the T7 promoter was induced by addition of arabinose, as described in Example 1. A protein of the expected molecular weight of about 31 kDa, as deduced from the nucleic acid sequence, was present in the induced culture, but was absent in the uninduced control.

Hydroxybutyryl-CoA dehydrogenase activity was determined by measuring the rate of oxidation of NADH as measured by the decrease in absorbance at 340 nm. A standard assay mixture contained 50 mM MOPS, pH 7.0, 1 mM DTT and 0.2 mM NADH. The cocktail was equilibrated for 5 min at 37° C. and then the cell free extract was added. Reactions were initiated by addition of the substrate, 0.1 mM acetoacetyl-CoA. In one typical assay, the specific activity of

30

the BHBD protein in the induced culture was determined to be 57.4 $\mu\text{mol mg}^{-1} \text{ min}^{-1}$ compared to 0.885 $\mu\text{mol mg}^{-1} \text{ min}^{-1}$ in the uninduced culture.

Example 4

Cloning and Expression of Crotonase

The purpose of this Example was to clone the crt gene from *C. acetobutylicum* (ATCC 824) and express it in *E. coli*. The crt gene was amplified from *C. acetobutylicum* (ATCC 824) genomic DNA using PCR.

The crt gene was cloned and expressed using the method described in Example 1. The crt gene was amplified from *C. acetobutylicum* (ATCC 824) genomic DNA by PCR using primers N3 and N4 (see Table 4), given as SEQ ID NOs:17 and 18, respectively, creating a 794 bp product. The forward primer incorporated four bases (CACC) immediately adjacent to the translational start codon to allow directional cloning into pENTR/SD/D-TOPO (Invitrogen) to generate the plasmid pENTRSDD-TOPOcrt. Clones were submitted for sequencing with M13 Forward and Reverse primers, given as SEQ ID NOs:45 and 46 respectively, to confirm that the genes inserted in the correct orientation and to confirm the sequence. The nucleotide sequence of the open reading frame (ORF) for this gene and its predicted amino acid sequence are given as SEQ ID NO:7 and SEQ ID NO:8, respectively.

To create an expression clone, the crt gene was transferred to the pDEST 14 (Invitrogen) vector by recombination to generate pDEST14crt. The pDEST14crt vector was transformed into BL21-AI cells and expression from the T7 promoter was induced by addition of arabinose, as described in Example 1. A protein of the expected molecular weight of about 28 kDa, as deduced from the nucleic acid sequence, was present in much greater amounts in the induced culture than in the uninduced control.

Crotonase activity was assayed as described by Stern (*Methods Enzymol.* 1, 559-566, (1954)). In one typical assay, the specific activity of the crotonase protein in the induced culture was determined to be 444 $\mu\text{mol mg}^{-1} \text{ min}^{-1}$ compared to 47 $\mu\text{mol mg}^{-1} \text{ min}^{-1}$ in the uninduced culture.

Example 5

Cloning and Expression of Butyryl-CoA Dehydrogenase

The purpose of this Example was to express the enzyme butyryl-CoA dehydrogenase, also referred to herein as trans-2-Enoyl-CoA reductase, in *E. coli*. The CAC0462 gene, a putative trans-2-enoyl-CoA reductase homolog, was cloned from *C. acetobutylicum* (ATCC 824) and expressed in *E. coli*. The CAC0462 gene was amplified from *C. acetobutylicum* (ATCC 824) genomic DNA using PCR.

The CAC0462 gene was cloned and expressed using the method described in Example 1. The CAC0462 gene was amplified from *C. acetobutylicum* (ATCC 824) genomic DNA by PCR using primers N17 and N21 (see Table 4), given as SEQ ID NOs:29 and 30, respectively, creating a 1.3 kbp product. The forward primer incorporated four bases (CACC) immediately adjacent to the translational start codon to allow directional cloning into pENTR/SD/D-TOPO (Invitrogen) to generate the plasmid pENTRSDD-TOPOCAC0462. Clones were submitted for sequencing with M13 Forward and Reverse primers, given as SEQ ID NO:45 and 46 respectively, to confirm that the genes inserted in the correct orientation and to confirm the sequence. Additional sequencing primers,

31

N22SeqF1 (SEQ ID NO:53), N22SeqF2 (SEQ ID NO:54), N22SeqF3 (SEQ ID NO:55), N23SeqR1 (SEQ ID NO:56), N23SeqR2 (SEQ ID NO:57), and N23SeqR3 (SEQ ID NO:58) (see Table 5) were needed to completely sequence the PCR product. The nucleotide sequence of the open reading frame (ORF) for this gene and the predicted amino acid sequence of the enzyme are given as SEQ ID NO:9 and SEQ ID NO:10, respectively.

To create an expression clone, the CAC0462 gene was transferred to the pDEST 14 (Invitrogen) vector by recombination to generate pDEST14CAC0462. The pDEST14CAC0462 vector was transformed into BL21-AI cells and expression from the T7 promoter was induced by addition of arabinose, as described in Example 1. Analysis by SDS-PAGE showed no overexpressed protein of the expected molecular weight in the negative control or in the induced culture. The *C. acetobutylicum* CAC0462 gene used many rare *E. coli* codons. To circumvent problems with codon usage the pRARE plasmid (Novagen) was transformed into BL21-AI cells harboring the pDEST14CAC0462 vector. Expression studies with arabinose induction were repeated with cultures carrying the pRARE vector. A protein of the expected molecular weight of about 46 kDa was present in the induced culture but not in the uninduced control.

Trans-2-enoyl-CoA reductase activity was assayed as described by Hoffmeister et al. (*J. Biol. Chem.* 280, 4329-4338 (2005)). In one typical assay, the specific activity of the TER CAC0462 protein in the induced culture was determined to be 0.694 $\mu\text{mol mg}^{-1} \text{min}^{-1}$ compared to 0.0128 $\mu\text{mol mg}^{-1} \text{min}^{-1}$ in the uninduced culture.

Example 6

Cloning and Expression of Butyraldehyde Dehydrogenase (Acetylating)

The purpose of this Example was to clone the ald gene from *C. beijerinckii* (ATCC 35702) and express it in *E. coli*. The ald gene was amplified from *C. beijerinckii* (ATCC 35702) genomic DNA using PCR.

The ald gene was cloned and expressed using the method described in Example 1. The ald gene was amplified from *C. beijerinckii* (ATCC 35702) genomic DNA (prepared from anaerobically grown cultures, as described in Example 1) by PCR using primers N27 F1 and N28 R1 (see Table 4), given as SEQ ID NOs:31 and 32 respectively, creating a 1.6 kbp product. The forward primer incorporated four bases (CACC) immediately adjacent to the translational start codon to allow directional cloning into pENTR/SD/D-TOPO (Invitrogen) to generate the plasmid pENTRSDD-TOPOald. Clones were submitted for sequencing with M13 Forward and Reverse primers, given as SEQ ID NOs:45 and 46 respectively, to confirm that the genes inserted in the correct orientation and to confirm the sequence. Additional sequencing primers, N31SeqF2 (SEQ ID NO:59), N31SeqF3 (SEQ ID NO:60), N31SeqF4 (SEQ ID NO:61), N32SeqR1 (SEQ ID NO:72), N31SeqR2 (SEQ ID NO:62), N31SeqR3 (SEQ ID NO:63), N31SeqR4 (SEQ ID NO:64), and N31SeqR5 (SEQ ID NO:65) (see Table 5) were needed to completely sequence the PCR product. The nucleotide sequence of the open reading frame (ORF) for this gene and the predicted amino acid sequence of the enzyme are given as SEQ ID NO:11 and SEQ ID NO:12, respectively.

To create an expression clone, the ald gene was transferred to the pDEST 14 (Invitrogen) vector by recombination to generate pDEST14ald. The pDEST14ald vector was transformed into BL21-AI cells and expression from the T7 pro-

32

moter was induced by addition of arabinose, as described in Example 1. A protein of the expected molecular weight of about 51 kDa, as deduced from the nucleic acid sequence, was present in the induced culture, but not in the uninduced control.

Acylating aldehyde dehydrogenase activity was determined by monitoring the formation of NADH, as measured by the increase in absorbance at 340 nm, as described by Husemann et al. (*Appl. Microbiol. Biotechnol.* 31:435-444 (1989)). In one typical assay, the specific activity of the Ald protein in the induced culture was determined to be 0.106 $\mu\text{mol mg}^{-1} \text{min}^{-1}$ compared to 0.01 $\mu\text{mol mg}^{-1} \text{min}^{-1}$ in the uninduced culture.

Example 7

Cloning and Expression of Butanol Dehydrogenase

The purpose of this Example was to clone the bdhB gene from *C. acetobutylicum* (ATCC 824) and express it in *E. coli*. The bdhB gene was amplified from *C. acetobutylicum* (ATCC 824) genomic DNA using PCR.

The bdhB gene was cloned and expressed using the method described in Example 1. The bdhB gene was amplified from *C. acetobutylicum* (ATCC 824) genomic DNA by PCR using primers N11 and N12 (see Table 4), given as SEQ ID NOs:25 and 26, respectively, creating a 1.2 kbp product. The forward primer incorporated four bases (CACC) immediately adjacent to the translational start codon to allow directional cloning into pENTR/SD/D-TOPO (Invitrogen) to generate the plasmid pENTRSDD-TOPObdhB. The translational start codon was also changed from "GTG" to "ATG" by the primer sequence. Clones were submitted for sequencing with M13 Forward and Reverse primers, given as SEQ ID NOs:45 and 46 respectively, to confirm that the genes inserted in the correct orientation and to confirm the sequence. Additional sequencing primers, N1SeqF1 (SEQ ID NO:66), N1SeqF2 (SEQ ID NO:67), N12SeqR1 (SEQ ID NO:68), and N12SeqR2 (SEQ ID NO:69), (see Table 5) were needed to completely sequence the PCR product. The nucleotide sequence of the open reading frame (ORF) for this gene and the predicted amino acid sequence of the enzyme are given as SEQ ID NO:13 and SEQ ID NO:14, respectively.

To create an expression clone, the bdhB gene was transferred to the pDEST 14 (Invitrogen) vector by recombination to generate pDEST14bdhB. The pDEST14bdhB vector was transformed into BL21-AI cells and expression from the T7 promoter was induced by addition of arabinose, as described in Example 1. A protein of the expected molecular weight of about 43 kDa, as deduced from the nucleic acid sequence, was present in the induced culture, but not in the uninduced control.

Butanol dehydrogenase activity was determined from the rate of oxidation of NADH as measured by the decrease in absorbance at 340 nm as described by Husemann and Papoutsakis, supra. In one typical assay, the specific activity of the BdhB protein in the induced culture was determined to be 0.169 $\mu\text{mol mg}^{-1} \text{min}^{-1}$ compared to 0.022 $\mu\text{mol mg}^{-1} \text{min}^{-1}$ in the uninduced culture.

Example 8

Cloning and Expression of Butanol Dehydrogenase

The purpose of this Example was to clone the bdhA gene from *C. acetobutylicum* 824 and express it in *E. coli*. The bdhA gene was amplified from *C. acetobutylicum* 824 genomic DNA using PCR.

The *bdhA* gene was cloned and expressed using the method described in Example 1. The *bdhA* gene was amplified from *C. acetobutylicum* 824 genomic DNA by PCR using primers N9 and N10 (see Table 4), given as SEQ ID NOs:23 and 24, respectively, creating a 1.2 kbp product. The forward primer incorporated four bases (CACC) immediately adjacent to the translational start codon to allow directional cloning into pENTR/SD/D-TOPO (Invitrogen) to generate the plasmid pENTRSDD-TOPO*bdhA*. Clones, given as SEQ ID NOs:45 and 46 respectively, to confirm that the genes inserted in the correct orientation and to confirm the sequence. Additional sequencing primers, N9SeqF1 (SEQ ID NO:70) and N10SeqR1 (SEQ ID NO:71), (see Table 5) were needed to completely sequence the PCR product. The nucleotide sequence of the open reading frame (ORF) for this gene and the predicted amino acid sequence of the enzyme are given as SEQ ID NO:15 and SEQ ID NO:16, respectively.

To create an expression clone, the *bdhA* gene was transferred to the pDEST 14 (Invitrogen) vector by recombination to generate pDEST14*bdhA*. The pDEST14*bdhA* vector was transformed into BL21-AI cells and expression from the T7 promoter was induced by addition of arabinose, as described in Example 1. A protein of the expected molecular weight of about 43 kDa, as deduced from the nucleic acid sequence, was present in the induced culture, but not in the uninduced control.

Butanol dehydrogenase activity was determined from the rate of oxidation of NADH as measured by the decrease in absorbance at 340 nm, as described by Husemann and Papoutsakis, supra. In one typical assay, the specific activity of the BdhA protein in the induced culture was determined to be 0.102 $\mu\text{mol mg}^{-1} \text{min}^{-1}$ compared to 0.028 $\mu\text{mol mg}^{-1} \text{min}^{-1}$ in the uninduced culture

Example 9

Construction of a Transformation Vector for the Genes in the 1-butanol Biosynthetic Pathway

Lower Pathway

To construct a transformation vector comprising the genes encoding the six steps in the 1-butanol biosynthetic pathway, the genes encoding the 6 steps in the pathway were divided into two operons. The upper pathway comprises the first four steps catalyzed by acetyl-CoA acetyltransferase, 3-hydroxybutyryl-CoA dehydrogenase, crotonase, and butyryl-CoA dehydrogenase. The lower pathway comprises the last two steps, catalyzed by butyraldehyde dehydrogenase and butanol dehydrogenase.

The purpose of this Example was to construct the lower pathway operon. Construction of the upper pathway operon is described in Example 10.

The individual genes were amplified by PCR with primers that incorporated restriction sites for later cloning and the forward primers contained an optimized *E. coli* ribosome binding site (AAAGGAGG). PCR products were TOPO cloned into the pCR 4Blunt-TOPO vector and transformed into *E. coli* Top10 cells (Invitrogen). Plasmid DNA was prepared from the TOPO clones and the sequence of the genes was verified. Restriction enzymes and T4 DNA ligase (New England Biolabs, Beverly, Mass.) were used according to manufacturer's recommendations. For cloning experiments,

restriction fragments were purified by gel electrophoresis using QIAquick Gel Extraction kit (Qiagen).

After confirmation of the sequence, the genes were subcloned into a modified pUC19 vector as a cloning platform. The pUC19 vector was modified by a HindIII/SapI digest, creating pUC19dHS. The digest removed the lac promoter adjacent to the MCS (multiple cloning site), preventing transcription of the operons in the vector.

The *ald* gene was amplified from *C. beijerinckii* ATCC 35702 genomic DNA by PCR using primers N58 and N59 (see Table 4), given as SEQ ID NOs:41 and 42, respectively, creating a 1.5 kbp product. The forward primer incorporated the restriction sites *Ava*I and *Bst*EII and a RBS (ribosome binding site). The reverse primer incorporated the *Hpa*I restriction site. The PCR product was cloned into pCRBlunt II-TOPO creating pCRBluntII-*ald*. Plasmid DNA was prepared from the TOPO clones and the sequence of the genes verified with primers M13 Forward (SEQ ID NO:45), M13 Reverse (SEQ ID NO:46), N31SeqF2 (SEQ ID NO:59), N31SeqF3 (SEQ ID NO:60), N31SeqF4 (SEQ ID NO:61), N32SeqR1 (SEQ ID NO:72), N31SeqR2 (SEQ ID NO:62), N31SeqR3 (SEQ ID NO:63), N31SeqR4 (SEQ ID NO:64), and N31SeqR5 (SEQ ID NO:65) (see Table 5).

The *bdhB* gene was amplified from *C. acetobutylicum* (ATCC 824) genomic DNA by PCR using primers N64 and N65 (see Table 4), given as SEQ ID NOs:43 and 44, respectively, creating a 1.2 kbp product. The forward primer incorporated an *Hpa*I restriction site and a RBS. The reverse primer incorporated a *Pme*I and a *Sph*I restriction site. The PCR product was cloned into pCRBlunt II-TOPO creating pCRBluntII-*bdhB*. Plasmid DNA was prepared from the TOPO clones and the sequence of the genes verified with primers M13 Forward (SEQ ID NO:45), M13 Reverse (SEQ ID NO:46), N11SeqF1 (SEQ ID NO:66), N11SeqF2 (SEQ ID NO:67), N12SeqR1 (SEQ ID NO:68), and N12SeqR2 (SEQ ID NO:69) (see Table 5).

To construct the lower pathway operon, a 1.2 kbp *Sph*I and *Hpa*I fragment from pCRBluntII-*bdhB*, a 1.4 kbp *Hpa*I and *Sph*I fragment from pCRBluntII-*ald*, and the large fragment from a *Ava*I and *Sph*I digest of pUC19dHS were ligated together. The three-way ligation created pUC19dHS-*ald-bdhB*.

The pUC19dHS-*ald-bdhB* vector was digested with *Bst*EII and *Pme*I releasing a 2.6 kbp fragment that was cloned into pBenBP, an *E. coli-Bacillus subtilis* shuttle vector. Plasmid pBenBP was created by modification of the pBE93 vector, which is described by Nagarajan, WO 93/24631 (Example 4). The *Bacillus amyloliquefaciens* neutral protease promoter (NPR), signal sequence and the *phoA* gene were removed from pBE93 with a *Nco*I/*Hind*III digest. The NPR promoter was PCR amplified from pBE93 by primers BenF and BenBPR, given by SEQ ID NOs:73 and 75, respectively. Primer BenBPR incorporated *Bst*EII, *Pme*I and *Hind*III sites downstream of the promoter. The PCR product was digested with *Nco*I and *Hind*III and the fragment was cloned into the corresponding sites in the vector pBE93 to create pBenBP. The lower operon fragment was subcloned into the *Bst*EII and *Pme*I sites in pBenBP creating pBen-*ald-bdhB*.

Assays for butyraldehyde dehydrogenase and butanol dehydrogenase activity were conducted on crude extracts using the methods described above. Both enzyme activities were demonstrated at levels above the control strain that contained an empty vector.

35

Example 10

Prophetic

Construction of a Transformation Vector for the Genes in the 1-butanol Biosynthetic Pathway

Upper Pathway

The purpose of this prophetic Example is to describe how to assemble the upper pathway operon. The general approach is the same as described in Example 9.

The *thlA* gene is amplified from *C. acetobutylicum* (ATCC 824) genomic DNA by PCR using primer pair N44 and N45 (see Table 4), given as SEQ ID NOs:33 and 34, respectively, creating a 1.2 kbp product. The forward primer incorporates a *SphI* restriction site and a ribosome binding site (RBS). The reverse primer incorporates *AscI* and *PstI* restriction sites. The PCR product is cloned into pCR4Blunt-TOPO creating pCR4Blunt-TOPO-*thlA*. Plasmid DNA is prepared from the TOPO clones and the sequence of the genes is verified with primers M13 Forward (SEQ ID NO:45), M13 Reverse (SEQ ID NO:46), N7SeqF1 (SEQ ID NO:47), and N7SeqR1 (SEQ ID NO:48) (see Table 5).

The *hbd* gene is amplified from *C. acetobutylicum* (ATCC 824) genomic DNA by PCR using primer pair N42 and N43 (see Table 4) given as SEQ ID NOs:35 and 36, respectively, creating a 0.9 kbp product. The forward primer incorporates a *Sall* restriction site and a RBS. The reverse primer incorporates a *SphI* restriction site. The PCR product is cloned into pCR4Blunt-TOPO creating pCR4Blunt-TOPO-*hbd*. Plasmid DNA is prepared from the TOPO clones and the sequence of the genes verified with primers M13 Forward (SEQ ID NO:45), M13 Reverse (SEQ ID NO:46), N5SeqF2 (SEQ ID NO:51), and N6SeqR2 (SEQ ID NO:52) (see Table 5).

The CAC0462 gene is codon optimized for expression in *E. coli* as primary host and *B. subtilis* as a secondary host. The new gene called CaTER, given as SEQ ID NO:76, is synthesized by Genscript Corp (Piscataway, N.J.). The gene CaTER is cloned in the pUC57 vector as a *BamHI*-*Sall* fragment and includes a RBS, producing plasmid pUC57-CaTER.

The *crt* gene is amplified from *C. acetobutylicum* (ATCC 824) genomic DNA by PCR using primer pair N38 and N39 (see Table 4), given as SEQ ID NOs:39 and 40, respectively, creating a 834 bp product. The forward primer incorporates *EcoRI* and *MluI* restriction sites and a RBS. The reverse primer incorporates a *BamHI* restriction site. The PCR product is cloned into pCR4Blunt-TOPO creating pCR4Blunt-TOPO-*crt*. Plasmid DNA is prepared from the TOPO clones and the sequence of the genes is verified with primers M13 Forward (SEQ ID NO:45) and M13 Reverse (SEQ ID NO:46) (see Table 5).

After confirmation of the sequence, the genes are subcloned into a modified pUC19 vector as a cloning platform. The pUC19 vector was modified by a *SphI*/*SapI* digest, creating pUC19dSS. The digest removed the *lac* promoter adjacent to the MCS, preventing transcription of the operons in the vector.

To construct the upper pathway operon pCR4Blunt-TOPO-*crt* is digested with *EcoRI* and *BamHI* releasing a 0.8 kbp *crt* fragment. The pUC19dSS vector is also digested with *EcoRI* and *BamHI* releasing a 2.0 kbp vector fragment. The *crt* fragment and the vector fragment are ligated together using T4 DNA ligase (New England Biolabs) to form pUC19dSS-*crt*. The CaTER gene is inserted into pUC19dSS-*crt* by digesting pUC57-CaTER with *BamHI* and *Sall*, releasing a 1.2 kbp CaTER fragment. The pUC19dSS-*crt* is

36

digested with *BamHI* and *Sall* and the large vector fragment is ligated with the CaTER fragment, creating pUC19dSS-*crt*-CaTER. To complete the operon a 884 bp *Sall* and *SphI* fragment from pCR4Blunt-TOPO-*hbd*, a 1.2 kb *SphI* and *PstI* *thlA* fragment from pCR4 Blunt-TOPO-*thlA* and the large fragment from a *Sall* and *PstI* digest of pUC19dSS-*crt*-CaTER are ligated. The product of the 3-way ligation is pUC19dSS-*crt*-CaTER-*hbd*-*thlA*.

The pUC19dSS-*crt*-CaTER-*hbd*-*thlA* vector is digested with *MluI* and *AscI* releasing a 4.1 kbp fragment that is cloned into a derivative of pBE93 (Caimi, WO2004/018645, pp. 39-40) an *E. coli*-*B. subtilis* shuttle vector, referred to as pBenMA. Plasmid pBenMA was created by modification of the pBE93 vector. The *Bacillus amyloliquefaciens* neutral protease promoter (NPR), signal sequence and the *phoA* gene are removed from pBE93 with a *NcoI*/*HindIII* digest. The NPR promoter is PCR amplified from pBE93 by primers BenF and BenMAR, given as SEQ ID NOS:73 and 74, respectively. Primer BenMAR incorporates *MluI*, *AscI*, and *HindIII* sites downstream of the promoter. The PCR product was digested with *NcoI* and *HindIII* and the fragment is cloned into the corresponding sites in the vector pBE93, creating pBenMA. The upper operon fragment is subcloned into the *MluI* and *AscI* sites in pBenMA creating pBen-*crt*-*hbd*-CaTER-*thlA*.

Example 11

Prophetic

Expression of the 1-butanol Biosynthetic Pathway in *E. coli*

The purpose of this prophetic Example is to describe how to express the 1-butanol biosynthetic pathway in *E. coli*.

The plasmids pBen-*crt*-*hbd*-CaTER-*thlA* and pBen-*ald*-*bdhB*, constructed as described in Examples 10 and 9, respectively, are transformed into *E. coli* NM522 (ATCC 47000) and expression of the genes in each operon is monitored by SDS-PAGE analysis, enzyme assay and Western analysis. For Westerns, antibodies are raised to synthetic peptides by Sigma-Genosys (The Woodlands, Tex.). After confirmation of expression of all the genes, pBen-*ald*-*bdhB* is digested with *EcoRI* and *PmeI* to release the NPR promoter-*ald*-*bdhB* fragment. The *EcoRI* digest of the fragment is blunt ended using the Klenow fragment of DNA polymerase (New England Biolabs, catalog no. M0210S). The plasmid pBen-*crt*-*hbd*-CaTER-*thlA* is digested with *PvuII* to create a linearized blunt ended vector fragment. The vector and NPR-*ald*-*bdhB* fragment are ligated, creating p1B1 O.1 and p1B1 O.2, containing the complete 1-butanol biosynthetic pathway with the NPR promoter-*ald*-*bdhB* fragment in opposite orientations. The plasmids p1B1 O.1 and p1B1 O.2 are transformed into *E. coli* NM522 and expression of the genes are monitored as previously described.

E. coli strain NM522/p1B1 O.1 or NM522/p1B1 O.1 is inoculated into a 250 mL shake flask containing 50 mL of medium and shaken at 250 rpm and 35° C. The medium is composed of: dextrose, 5 g/L; MOPS, 0.05 M; ammonium sulfate, 0.01 M; potassium phosphate, monobasic, 0.005 M; S10 metal mix, 1% (v/v); yeast extract, 0.1% (w/v); casamino acids, 0.1% (w/v); thiamine, 0.1 mg/L; proline, 0.05 mg/L; and biotin 0.002 mg/L, and is titrated to pH 7.0 with KOH. S10 metal mix contains: MgCl₂, 200 mM; CaCl₂, 70 mM; MnCl₂, 5 mM; FeCl₃, 0.1 mM; ZnCl₂, 0.1 mM; thiamine hydrochloride, 0.2 mM; CuSO₄, 172 μM; CoCl₂, 253 μM;

and Na₂MoO₄, 242 μM. After 18 to 24 h, 1-butanol is detected by HPLC or GC analysis, as described in the General Methods section.

Example 12

Prophetic

Expression of the 1-butanol Biosynthetic Pathway in *Bacillus subtilis*

The purpose of this prophetic Example is to describe how to express the 1-butanol biosynthetic pathway in *Bacillus subtilis*. The same approach as described in Example 11 is used.

The upper and lower operons constructed as described in Examples 10 and 9, respectively, are used. The plasmids p1B1 O.1 and p1B1 O.2 are transformed into *Bacillus subtilis* BE1010 (*J. Bacteriol.* 173:2278-2282 (1991)) and expression of the genes in each operon is monitored as described in Example 11.

B. subtilis strain BE1010/p1B1 O.1 or BE11010/p1B1 O.2 is inoculated into a 250 mL shake flask containing 50 mL of medium and shaken at 250 rpm and 35° C. for 18 h. The medium is composed of: dextrose, 5 g/L; MOPS, 0.05 M; glutamic acid, 0.02 M; ammonium sulfate, 0.01 M; potassium phosphate, monobasic buffer, 0.005 M; S10 metal mix (as described in Example 11), 1% (v/v); yeast extract, 0.1% (w/v); casamino acids, 0.1% (w/v); tryptophan, 50 mg/L; methionine, 50 mg/L; and lysine, 50 mg/L, and is titrated to pH 7.0 with KOH. After 18 to 24 h, 1-butanol is detected by HPLC or GC analysis, as described in the General Methods section.

Example 13

Production of 1-butanol from Glucose using Recombinant *E. coli*

This Example describes the production of 1-butanol in *E. coli*. Expression of the genes encoding the 6 steps of the 1-butanol biosynthetic pathway was divided into three operons. The upper pathway comprised the first four steps encoded by thlA, hbd, crt and EgTER in one operon. The next step, encoded by ald, was provided by a second operon. The last step in the pathway, encoded by yqhD, was provided in a third operon. 1-butanol production was demonstrated in *E. coli* strains comprising the three operons.

Unless otherwise indicated in the text, cloning primers described in this Example are referenced by their SEQ ID NO: in Table 4, and sequencing and PCR screening primers are referenced by their SEQ ID NO: in Table 5.

Acetyl-CoA acetyltransferase.

The thlA gene was amplified from *C. acetobutylicum* (ATCC 824) genomic DNA by PCR using primer pair N44 and N45 (see Table 4), given as SEQ ID NOs:33 and 34, respectively, creating a 1.2 kbp product. The forward primer incorporated a SphI restriction site and a ribosome binding site (RBS). The reverse primer incorporated AscI and PstI restriction sites. The PCR product was cloned into pCR4Blunt-TOPO (Invitrogen Corp., Carlsbad, Calif.) creating pCR4Blunt-TOPO-thlA. Plasmid DNA was prepared from the TOPO clones and the sequence of the genes was verified with primers M13 Forward (SEQ ID NO:45), M13 Reverse (SEQ ID NO:46), N7SeqF1 (SEQ ID NO:47), and N7SeqR1 (SEQ ID NO:48) (see Table 5).

3-hydroxybutyryl-CoA dehydrogenase

The hbd gene was amplified from *C. acetobutylicum* (ATCC 824) genomic DNA by PCR using primer pair N42 and N43 (see Table 4) given as SEQ ID NOs:35 and 36, respectively, creating a 0.9 kbp product. The forward primer incorporated a SalI restriction site and a RBS. The reverse primer incorporated a SphI restriction site. The PCR product was cloned into pCR4Blunt-TOPO creating pCR4Blunt-TOPO-hbd. Plasmid DNA was prepared from the TOPO clones and the sequence of the genes verified with primers M13 Forward (SEQ ID NO:45), M13 Reverse (SEQ ID NO:46), N5SeqF2 (SEQ ID NO:51), and N6SeqR2 (SEQ ID NO:52) (see Table 5).

Crotonase.

The crt gene was amplified from *C. acetobutylicum* (ATCC 824) genomic DNA by PCR using primer pair N38 and N39 (see Table 4), given as SEQ ID NOs:39 and 40, respectively, creating a 834 bp product. The forward primer incorporated EcoRI and MluI restriction sites and a RBS. The reverse primer incorporated a BamHI restriction site. The PCR product was cloned into pCR4Blunt-TOPO creating pCR4Blunt-TOPO-crt. Plasmid DNA was prepared from the TOPO clones and the sequence of the genes was verified with primers M13 Forward (SEQ ID NO:45) and M13 Reverse (SEQ ID NO:46) (see Table 5).

Butyryl-CoA Dehydrogenase (trans-2-enoyl-CoA reductase).

The CAC0462 gene was synthesized for enhanced codon usage in *E. coli* as primary host and *B. subtilis* as a secondary host. The new gene (CaTER, SEQ ID NO:76) was synthesized and cloned by Genscript Corporation (Piscataway, N.J.) in the pUC57 vector as a BamHI-SalI fragment and includes a RBS.

An alternative gene for butyryl-CoA dehydrogenase from *Euglena gracilis* (TER, GenBank No. Q5EU90) was synthesized for enhanced codon usage in *E. coli* and *Bacillus subtilis*. The gene was synthesized and cloned by GenScript Corporation into pUC57 creating pUC57::EgTER. Primers N85 and N86, (SEQ ID NO: 80 and 81 respectively) together with pUC57::EgTER as template DNA, provided a PCR fragment comprising 1224 bp from pUC57::EgTER DNA. The sequence of the 1224 bp is given as SEQ ID NO:77, where bp 1-1218 is the coding sequence (cds) of EgTER (opt). EgTER (opt) is a codon optimized TER gene, lacking the normal mitochondrial presequence so as to be functional in *E. coli* (Hoffmeister et al., *J. Biol. Chem.* 280:4329 (2005)).

EgTER(opt) was cloned into pCR4Blunt-TOPO and its sequence was confirmed with primers M13 Forward (SEQ ID NO:45) and M13 Reverse (SEQ ID NO:46). Additional sequencing primers N62SeqF2 (SEQ ID NO:114), N62SeqF3 (SEQ ID NO:115), N62SeqF4 (SEQ ID NO:116), N63SeqR1 (SEQ ID NO:117), N63SeqR2 (SEQ ID NO:118), N63SeqR3 (SEQ ID NO:119) and N63SeqR4 (SEQ ID NO:120) were needed to completely sequence the PCR product. The 1.2 kbp EgTER(opt) sequence was then excised with HincII and PmeI and cloned into pET23+ (Novagen) linearized with HincII. Orientation of the EgTER(opt) gene to the promoter was confirmed by colony PCR screening with primers T7Primer and N63SeqR2 (SEQ ID NOs:82 and 118 respectively). The resulting plasmid, pET23+::EgTER(opt), was transformed into BL21 (DE3) (Novagen) for expression studies.

Trans-2-enoyl-CoA reductase activity was assayed as described by Hoffmeister et al., *J. Biol. Chem.* 280:4329 (2005). In a typical assay, the specific activity of the EgTER (opt) protein in the induced BL21 (DE3)/pET23+::EgTER

(opt) culture was determined to be $1.9 \mu\text{mol mg}^{-1} \text{min}^{-1}$ compared to $0.547 \mu\text{mol mg}^{-1} \text{min}^{-1}$ in the uninduced culture.

The EgTER(opt) gene was then cloned into the pTrc99a vector under the control of the trc promoter. The EgTER(opt) gene was isolated as a 1287-bp BamHI/SalI fragment from pET23+::EgTER(opt). The 4.2 kbp vector pTrc99a was linearized with BamHI/SalI. The vector and fragment were ligated creating the 5.4 kbp pTrc99a-EgTER(opt). Positive clones were confirmed by colony PCR with primers Trc99aF and N63SeqR3 (SEQ ID NOs:83 and 119 respectively) producing a 0.5 kb product.

Construction of Plasmid pTrc99a-E-C-H-T Comprising Genes Encoding acetyl-CoA acetyltransferase (thlA), 3-hydroxybutyryl-CoA dehydrogenase (hbd), crotonase (crt), and butyryl-CoA dehydrogenase (trans-2-enoyl-CoA reductase. EgTER(opt))

To initiate the construction of a four gene operon comprising the upper pathway (EgTER(opt), crt, hbd and thlA), pCR4Blunt-TOPO-crt was digested with EcoRI and BamHI releasing a 0.8 kbp crt fragment. The pUC19dSS vector (described in Example 10) was also digested with EcoRI and BamHI releasing a 2.0 kbp vector fragment. The crt fragment and the vector fragment were ligated together using T4 DNA ligase (New England Biolabs) to form pUC19dSS-crt. The CaTER gene was inserted into pUC19dSS-crt by digesting pUC57-CaTER with BamHI and SalI, releasing a 1.2 kbp CaTER fragment. The pUC19dSS-crt was digested with BamHI and SalI and the large vector fragment was ligated with the CaTER fragment, creating pUC19dSS-crt-CaTER. To complete the operon a 884 bp SalI and SphI fragment from pCR4Blunt-TOPO-hbd, a 1.2 kb SphI and PstI thlA fragment from pCR4Blunt-TOPO-thlA and the large fragment from a SalI and PstI digest of pUC19dSS-crt-CaTER were ligated. The product of the 3-way ligation was named pUC19dSS-crt-CaTER-hbd-thlA or pUC19dss::Operon1.

Higher butyryl-CoA dehydrogenase activity was obtained from pTrc99a-EgTER(opt) than from CaTER constructs, thus, an operon derived from pTrc99a-EgTER(opt) was constructed. The CaTER gene was removed from pUC19dss::Operon1 by digesting with BamHI/SalI and gel purifying the 5327-bp vector fragment. The vector was treated with Klenow and religated creating pUC19dss::Operon1 Δ CaTER. The 2934-bp crt-hbd-thlA (C-H-T) fragment was then isolated as a EcoRI/PstI fragment from pUC19dss::Operon1 Δ CaTER. The C-H-T fragment was treated with Klenow to blunt the ends. The vector pTrc99a-EgTER(opt) was digested with SalI and the ends treated with Klenow. The blunt-ended vector and the blunt-ended C-H-T fragment were ligated to create pTrc99a-E-C-H-T. Colony PCR reactions were performed with primers N62SeqF4 and N5SeqF4 (SEQ ID NOs: 116 and 84 respectively) to confirm the orientation of the insert.

Construction of Plasmids pBHR T7-ald and pBHR-Ptrc-ald(opt) Comprising Genes Encoding butyraldehyde dehydrogenase (ald and ald(opt)).

The PT7-ald operon was sub-cloned from pDEST14-ald (Example 6) into the broad host range plasmid pBHR1 (MoBitec, Goettingen, Germany) to create pBHR1PT7-ald. The pBHR1 plasmid is compatible with pUC19 or pBR322 plasmids so pBHR1PT7-ald can be used in combination with pUC19 or pBR322 derivatives carrying the upper pathway operon for 1-butanol production in *E. coli*. The pDEST14-ald plasmid was digested with Bgl II and treated with the Klenow fragment of DNA polymerase to make blunt ends. The plasmid was then digested with EcoRI and the 2,245 bp PT7-ald fragment was gel-purified. Plasmid pBHR1 was digested with ScaI and EcoRI and the 4,883 bp fragment was gel-purified. The PT7-ald fragment was ligated with the pBHR1

vector, creating pBHR T7-ald. Colony PCR amplification of transformants with primers T-ald(BamHI) and B-ald (EgTER) (SEQ ID NOs:85 and 86 respectively) confirmed the expected 1.4 kb PCR product. Restriction mapping of pBHR T7-ald clones with EcoRI and DrdI confirmed the expected 4,757 and 2,405 bp fragments.

For butyraldehyde dehydrogenase activity assays, the plasmid pBHR T7-ald was transformed into BL21Star™ (DE3) cells (Invitrogen) and expression from the T7 promoter was induced by addition of L-arabinose as described in Example 1. Acylating aldehyde dehydrogenase activity was determined by monitoring the formation of NADH, as measured by the increase in absorbance at 340 nm, as described in Example 6.

An alternative DNA sequence for the ald gene from *Clostridium beijerinckii* ATCC 35702 was synthesized (optimizing for codon usage in *E. coli* and *Bacillus subtilis*) and cloned into pUC57 by GenScript Corporation (Piscataway, N.J.), creating the plasmid pUC57-ald(opt). pUC57-ald(opt) was digested with SacI and SalI to release a 1498 bp fragment comprising the codon optimized gene, aid(opt) and a RBS already for *E. coli*. The sequence of the 1498 bp fragment is given as SEQ ID NO:78.

pTrc99a was digested with SacI and SalI giving a 4153 bp vector fragment, which was ligated with the 1498 bp aid(opt) fragment to create pTrc-ald(opt). Expression of the synthetic gene, aid(opt), is under the control of the IPTG-inducible Ptrc promoter.

The Ptrc-aid(opt) operon was subcloned into the broad host range plasmid pBHR1 (MoBitec) in order to be compatible with the upper pathway plasmid described above. The Ptrc-aid(opt) fragment was PCR-amplified from pTrc99a::ald(opt) with T-Ptrc(BspEI) and B-ald(opt)(ScaI), (SEQ ID NOs: 87 and 88 respectively) incorporating BspEI and ScaI restriction sites within the corresponding primers. The PCR product was digested with BspEI and ScaI. The plasmid pBHR1 was digested with ScaI and BspEI and the 4,883 bp fragment was gel-purified. The Ptrc-aid(opt) fragment was ligated with the pBHR1 vector, creating pBHR-PcatPtrc-ald(opt). Restriction mapping of the pBHR-PcatPtrc-ald(opt) clones with ScaI and BspEI confirmed the expected 4,883 and 1,704 bp fragments. To remove the plasmid-born cat promoter (Pcat) region, plasmid pBHR-PcatPtrc-ald(opt) was digested with BspEI and AatII and the 6,172 bp fragment was gel-purified. T-BspEIAatII and B-BspEIAatII (SEQ ID NOs: 89 and 90 respectively) were mixed in a solution containing 50 mM NaCl, 10 mM Tris-HCl, and 10 mM MgCl₂ (pH7.9) to a final concentration of 100 μ M and hybridized by incubating at 75° C. for 5 min and slowly cooling to room temperature. The hybridized oligonucleotides were ligated with the 6,172 bp fragment, creating pBHR-Ptrc-ald(opt).

Construction of *E. coli* Strains Expressing Butanol Dehydrogenase (yqhD).

E. coli contains a native gene (yqhD) that was identified as a 1,3-propanediol dehydrogenase (U.S. Pat. No. 6,514,733). The yqhD gene has 40% identity to the gene adhB in *Clostridium*, a probable NADH-dependent butanol dehydrogenase. The yqhD gene was placed under the constitutive expression of a variant of the glucose isomerase promoter 1.6GI (SEQ ID NO:91) in *E. coli* strain MG1655 1.6yqhD::Cm (WO 2004/033646) using λ Red technology (Datsenko and Wanner, *Proc. Natl. Acad. Sci. U.S.A.* 97:6640 (2000)). Similarly, the native promoter was replaced by the 1.5GI promoter (WO 2003/089621) (SEQ ID NO:92), creating strain MG1655 1.5GI-yqhD::Cm, thus, replacing the 1.6GI promoter of MG1655 1.6yqhD::Cm with the 1.5GI promoter.

41

A P1 lysate was prepared from MG1655 1.5GI-yqhD::Cm and the cassette moved to expression strains, MG1655 (DE3), prepared from *E. coli* strain MG1655 and a lambda DE3 lysogenization kit (Invitrogen), and BL21 (DE3) (Invitrogen) creating MG1655 (DE3) 1.5GI-yqhD::Cm and BL21 (DE3) 1.5GI-yqhD::Cm, respectively.

Demonstration of 1-butanol Production from Recombinant *E. coli*.

E. coli strain MG1655 (DE3) 1.5GI-yqhD::Cm was transformed with plasmids pTrc99a-E-C-H-T and pBHR T7-ald to produce the strain, MG1655 (DE3) 1.5GI-yqhD::Cm/pTrc99a-E-C-H-T/pBHR T7-ald. Two independent isolates were initially grown in LB medium containing 50 µg/mL kanamycin and 100 µg/mL carbenicillin. The cells were used to inoculate shake flasks (approximately 175 mL total volume) containing 15, 50 and 150 mL of TM3a/glucose medium (with appropriate antibiotics) to represent high, medium and low oxygen conditions, respectively. TM3a/glucose medium contained (per liter): 10 g glucose, 13.6 g KH₂PO₄, 2.0 g citric acid monohydrate, 3.0 g (NH₄)₂SO₄, 2.0 g MgSO₄·7H₂O, 0.2 g CaCl₂·2H₂O, 0.33 g ferric ammonium citrate, 1.0 mg thiamine.HCl, 0.50 g yeast extract, and 10 mL trace elements solution, adjusted to pH 6.8 with NH₄OH. The solution of trace elements contained: citric acid.H₂O (4.0 g/L), MnSO₄·H₂O (3.0 g/L), NaCl (1.0 g/L), FeSO₄·7H₂O (0.10 g/L), CoCl₂·6H₂O (0.10 g/L), ZnSO₄·7H₂O (0.10 g/L), CuSO₄·5H₂O (0.010 g/L), H₃BO₃ (0.010 g/L), and Na₂MoO₄·2H₂O (0.010 g/L). The flasks were inoculated at a starting OD₆₀₀ of ≤0.01 units and incubated at 34° C. with shaking at 300 rpm. The flasks containing 15 and 50 mL of medium were capped with vented caps; the flasks containing 150 mL, were capped with non-vented caps to minimize air exchange. IPTG was added to a final concentration of 0.04 mM; the OD₆₀₀ of the flasks at the time of addition was ≥0.4 units.

Approximately 15 h after induction, an aliquot of the broth was analyzed by HPLC (Shodex Sugar SH1011 column) with refractive index (RI) detection and GC (Varian CP-WAX 58(FFAP) CB column, 25 m×0.25 mm id×0.2 µm film thickness) with flame ionization detection (FID) for 1-butanol content, as described in the General Methods section. The results of the 1-butanol determinations are given in Table 6.

TABLE 6

Production of 1-butanol by <i>E. coli</i> strain MG1655 (DE3) 1.5GI-yqhD::Cm/pTrc99a-E-C-H-T/pBHR T7-ald.			
Strain	O ₂ Level	1-butanol, mM	molar yield, %
MG1655 a	high	0.11	0.2
MG1655 b	high	0.12	0.2
MG1655 a	medium	0.13	0.3
MG1655 b	medium	0.13	0.2
MG1655 a	low	0.15	0.4
MG1655 b	low	0.18	0.5

Values were determined from HPLC analysis.

Strain suffixes "a" and "b" indicate independent isolates.

The two independent isolates of MG1655 (DE3) 1.5GI-yqhD::Cm/pTrc99a-E-C-H-T/pBHR T7-ald were tested for 1-butanol production in an identical manner except that the medium contained 5 g/L yeast extract. The results are shown in Table 7.

42

TABLE 7

Production of 1-butanol by <i>E. coli</i> strain MG1655 (DE3) 1.5GI-yqhD::Cm/pTrc99a-E-C-H-T/pBHR T7-ald.			
Strain	O ₂ Level	1-butanol, mM	molar yield, %
MG1655 a	high	—	—
MG1655 b	high	—	—
MG1655 a	medium	0.08	0.1
MG1655 b	medium	0.06	0.1
MG1655 a	low	0.14	0.3
MG1655 b	low	0.14	0.3

Quantitative values were determined from HPLC analysis.

“—” = not detected.

Strain suffixes "a" and "b" indicate independent isolates.

E. coli strain BL21 (DE3) 1.5GI-yqhD::Cm was transformed with plasmids pTrc99a-E-C-H-T and pBHR T7-ald to produce the strain, BL21 (DE3) 1.5GI-yqhD::Cm/pTrc99a-E-C-H-T/pBHR T7-ald. Two independent isolates were tested for 1-butanol production exactly as described above. The results are given in Tables 8 and 9.

TABLE 8

Production of 1-butanol by <i>E. coli</i> strain BL21 (DE3) 1.5GI-yqhD::Cm/pTrc99a-E-C-H-T/pBHR T7-ald.			
Strain	O ₂ Level	1-butanol, mM	molar yield, %
DE a	high	+	+
DE b	high	—	—
DE a	medium	0.80	1.4
DE b	medium	0.77	1.4
DE a	low	0.06	0.2
DE b	low	0.07	0.2

Quantitative values were determined from HPLC analysis.

“—” indicates not detected.

“+” indicates positive, qualitative identification by GC with a lower detection limit than with HPLC.

Strain suffixes "a" and "b" indicate independent isolates.

TABLE 9

Production of 1-butanol by <i>E. coli</i> strain BL21 (DE3) 1.5GI-yqhD::Cm/pTrc99a-E-C-H-T/pBHR T7-ald.			
Strain	O ₂ Level	1-butanol, mM	molar yield, %
DE a	high	+	+
DE b	high	+	+
DE a	medium	0.92	1.7
DE b	medium	1.03	1.9
DE a	low	+	+
DE b	low	+	+

Quantitative values were determined from HPLC analysis.

“—” indicates not detected.

“+” indicates positive, qualitative identification by GC with a lower detection limit than with HPLC.

Strain suffixes "a" and "b" indicate independent isolates.

E. coli strain MG1655 1.5GI-yqhD::Cm was transformed with plasmids pTrc99a-E-C-H-T and pBHR-Ptrc-ald(opt) to produce the strain, MG1655 1.5GI-yqhD::Cm/pTrc99a-E-C-H-T/pBHR-Ptrc-ald(opt). Two isolates were initially grown in LB medium containing 50 µg/mL kanamycin and 100 µg/mL carbenicillin. The cells were used to inoculate shake flasks (approximately 175 mL total volume) containing 50 and 150 mL of TM3a/glucose medium (with appropriate antibiotics). The flasks were inoculated at a starting OD₅₅₀ of ≤0.04 units and incubated as described above, with and without induction. IPTG was added to a final concentration of 0.4 mM; the OD₅₅₀ of the flasks at the time of addition was between 0.6 and 1.2 units. In this case, induction was not

necessary for 1-butanol pathway gene expression because of the leakiness of the IPTG inducible promoters and the constitutive nature of the 1.5GI promoter; however, induction provided a wider range of expression.

Approximately 15 h after induction, an aliquot of the broth was analyzed by GC with flame ionization detection for 1-butanol content, as described above. The results are given in Table 10. For the recombinant *E. coli* strains, 1-butanol was produced in all cases; in separate experiments, wild type *E. coli* strains were shown to produce no detectable 1-butanol (data not shown).

TABLE 10

Production of 1-butanol by <i>E. coli</i> strain MG1655 1.5GI-yqhD::Cm/ pTtrc99a-E-C-H-T/pBHR-Ptrc-ald(opt).			
Strain	O ₂ Level	1-butanol, mM	IPTG Induction
MG1655 a	medium	0.14	No
MG 1655 b	medium	0.14	No
MG1655 a	medium	0.03	Yes
MG 1655 b	medium	0.07	Yes
MG1655 a	low	0.04	No
MG 1655 b	low	0.04	No
MG1655 a	low	0.02	Yes
MG 1655 b	low	0.03	Yes

Strain suffixes "a" and "b" indicate separate isolates.

Example 14

Production of 1-butanol from Glucose using Recombinant *B. subtilis*

This Example describes the production of 1-butanol in *Bacillus subtilis*. The six genes of the 1-biosynthetic pathway, encoding six enzyme activities, were split into two operons for expression. The first three genes of the pathway (thl, hbd, and crt) were integrated into the chromosome of *Bacillus subtilis* BE1010 (Payne and Jackson, J. Bacteriol. 173:2278-2282 (1991)). The last three genes (EgTER, aid, and bdhB) were cloned into an expression plasmid and transformed into the *Bacillus* strain carrying the integrated 1-butanol genes.

Unless otherwise indicated in the text, cloning primers described in this Example are referenced by their SEQ ID NO: in Table 4, and sequencing and PCR screening primers are referenced by their SEQ ID NO: in Table 5.

Integration Plasmid.

Plasmid pFP988 is a *Bacillus* integration vector that contains an *E. coli* replicon from pBR322, an ampicillin antibiotic marker for selection in *E. coli* and two sections of homology to the sacB gene in the *Bacillus* chromosome that directs integration of the vector and intervening sequence by homologous recombination. Between the sacB homology regions is the Pamy promoter and signal sequence that can direct the synthesis and secretion of a cloned gene, a His-Tag and erythromycin as a selectable marker for *Bacillus*. The Pamy promoter and signal sequence is from *Bacillus amyloliquefaciens* alpha-amylase. The promoter region also contains the lacO sequence for regulation of expression by a lacI repressor protein. The sequence of pFP988 (6509 bp) is given as SEQ ID NO:79.

Since the 1-butanol pathway genes were to be expressed in the cytoplasm, the amylase signal sequence was deleted. Plasmid pFP988 was amplified with primers Pamy/lacO F and Pamy/lacO R creating a 317 bp (0.3 kbp) product that contained the Pamy/lacO promoter. The 5' end of the Pamy/lacO F primer incorporated a BsrGI restriction site followed by an

EcoRI site. The 5' end of the Pamy/lacO R primer incorporated a BsrGI restriction site followed by a PmeI restriction site. The PCR product was TOPO cloned into pCR4Blunt-TOPO creating pCR4Blunt-TOPO-Pamy/lacO. Plasmid DNA was prepared from overnight cultures and submitted for sequencing with M13 Forward and M13 Reverse primers (SEQ ID NO:45 and SEQ ID NO:46, respectively) to ensure no mutation had been introduced into the promoter. A clone of pCR4Blunt-TOPO-Pamy/lacO was digested with BsrGI and the 0.3 kbp fragment was gel-purified. The vector pFP988 was digested with BsrGI resulting in deletion of 11 bp from the 5' sacB homology region and the removal of the Pamy/lacO promoter and signal sequence and His-tag. The 6 kbp BsrGI digested vector was gel-purified and ligated with Pamy/lacO BsrGI insert. The resulting plasmids were screened with primers Pamy SeqF2 and Pamy SeqR to determine orientation of the promoter. The correct clone restored the Pamy/lacO promoter to its original orientation and was named pFP988Dss.

The cassette with genes thl-crt was constructed by SOE (splicing by overlap extension). The genes were amplified using as template pUC19dss::Operon1. The thl primers were Top TF and Bot TR amplifying a 0.9 kbp product. The crt primers were Top CF and Bot CR amplifying a 1.3 kbp product. The two genes were joined by SOE with PCR amplification using primers Top TF and Bot CR generating a 2.1 kbp product that was TOPO cloned into pCR4Blunt-TOPO creating pCR4Blunt-TOPO-T-C. Clones were submitted for sequencing to confirm the sequence. The plasmid pCR4Blunt-TOPO-T-C was digested with BstEII and PmeI releasing a 2.1 kbp fragment that was gel-purified. The insert was treated with Klenow polymerase to blunt the BstEII site. Vector pFP988Dss was digested with PmeI and treated with calf intestinal alkaline phosphatase (New England BioLabs) to prevent self-ligation. The 2.1 kbp thl-crt fragment and the digested pFP988Dss were ligated and transformed into *E. coli* Top10 cells. Transformants were screened by PCR amplification with Pamy SeqF2 and N7SeqR2 for a 0.7 kbp product, the correct product was called pFP988Dss-T-C.

Construction of the thl-crt cassette created unique SalI and SpeI sites between the two genes. To add the hbd gene to the cassette, the hbd gene was subcloned from pCR4Blunt-TOPO-hbd as a 0.9 kbp SalI/SpeI fragment. Vector pFP988Dss-T-C was digested with SalI and SpeI and the 8 kbp vector fragment was gel-purified. The vector and hbd insert were ligated and transformed into *E. coli* Top10 cells. Transformants were screened by PCR amplification with primers Pamy SeqF and N3SeqF3 for a 3.0 kbp fragment. The resulting plasmid was named pFP988Dss-T-H-C.

The Pamy promoter subsequently was replaced with the Pspac promoter from plasmid pMUTIN4 (Vagner et al., *Microbiol.* 144:3097-3104 (1998)). The Pspac promoter was amplified from pMUTIN4 with primers Spac F and Spac R as a 0.4 kbp product and TOPO cloned into pCR4Blunt-TOPO. Transformants were screened by PCR amplification with M13 Forward and M13 Reverse primers for the presence of a 0.5 kbp insert. Positive clones were submitted for sequencing with the same primers. Plasmid pCR4Blunt-TOPO-Pspac was digested with SmaI and XhoI and the 0.3 kbp fragment was gel-purified. Vector pFP988Dss-T-H-C was digested with SmaI and XhoI and the 9 kbp vector was isolated by gel purification. The digested vector and Pspac insert were ligated and transformed into *E. coli* Top10 cells. Transformants were screened by PCR amplification with primers SpacF Seq and N7SeqR2. Positive clones gave a 0.7 kbp product. Plasmid DNA was prepared from positive clones and further screened by PCR amplification with primers SpacF

Seq and N3SeqF2. Positive clones gave a 3 kbp PCR product and were named pFP988DssPspac-T-H-C.

Integration into *B. subtilis* BE010 to Form *B. subtilis* Δ*sacB*::T-H-C::erm #28 comprising exogenous *thl*, *hbd*, and *crt* Genes

Competent cells of *B. subtilis* BE1010 were prepared as described in Doyle et al., *J. Bacteriol.* 144:957-966 (1980). Competent cells were harvested by centrifugation and the cell pellets were resuspended in a small volume of the cell supernatant. To 1 volume of competent cells, 2 volumes of SPII-EGTA medium (*Methods for General and Molecular Bacteriology*, P. Gerhardt et al., Eds, American Society for Microbiology, Washington, D.C. (1994)) was added. Aliquots of 0.3 mL of cells were dispensed into test tubes and the plasmid pFP988DssPspac-T-H-C was added to the tubes. Cells were incubated for 30 minutes at 37° C. with shaking, after which 0.1 mL of 10% yeast extract was added to each tube and the cells were further incubated for 60 min. Transformants were plated for selection on LB erythromycin plates using the double agar overlay method (*Methods for General and Molecular Bacteriology*, supra). Transformants were initially screened by PCR amplification with primers PamySeqF and N5SeqF3. Positive clones that amplified the expected 2 kbp PCR product were further screened by PCR amplification. If insertion of the cassette into the chromosome had occurred via a double crossover event then primer set *sacB* Up and N7SeqR2 and primer set *sacB* Dn and N4SeqR3 would amplify a 1.7 kbp and a 2.7 kbp product respectively. A positive clone was identified and named *B. subtilis* Δ*sacB*::T-H-C::erm #28.

Plasmid Expression of EgTER, aid, and bdhB genes.

The three remaining 1-butanol genes were expressed from plasmid pHT01 (MoBitec). Plasmid pHT01 is a *Bacillus-E. coli* shuttle vector that replicates via a theta mechanism. Cloned proteins are expressed from the GroEL promoter fused to a *laco* sequence. Downstream of the *laco* is the efficient RBS from the *gsiB* gene followed by a MCS. The aid gene was amplified by PCR with primers AF BamHI and AR Aat2 using pUC19dHS-ald-bdhB (described in Example 9) as template, creating a 1.4 kbp product. The product was TOPO cloned into pCR4-TOPO and transformed into *E. coli* Top10 cells. Transformants were screened with M13 Forward and M13 Reverse primers. Positive clones amplified a 1.6 kbp product. Clones were submitted for sequencing with primers M13 forward and M13 reverse, N31SeqF2, N31SeqF3, N32SeqR2, N32SeqR3 and N32SeqR4. The plasmid was named pCR4TOPO-B/A-ald.

Vector pHT01 and plasmid pCR4TOPO-B/A-ald were both digested with BamHI and AatII. The 7.9 kbp vector fragment and the 1.4 kbp aid fragment were ligated together to create pHT01-aid. The ligation was transformed into *E. coli* Top10 cells and transformants were screened by PCR amplification with primers N31 SeqF1 and HT R for a 1.3 kbp product.

To add the last two steps of the pathway to the pHT01 vector, two cloning schemes were designed. For both schemes, EgTER and bdhB were amplified together by SOE. Subsequently, the EgTER-bdh fragment was either cloned into pHT01-ald creating pHT01-ald-EB or cloned into pCR4-TOPO-B/A-ald creating pCR4-TOPO-ald-EB. The ald-EgTer-bdhB fragment from the TOPO vector was then cloned into pHT01 creating pHT01-AEB.

An EgTER-bdhB fragment was PCR amplified using primers Forward 1 (E) and Reverse 2 (B), using template DNA given as SEQ ID NO:208. The resulting 2.5 kbp PCR product was TOPO cloned into pCR4Blunt-TOPO, creating pCR4Blunt-TOPO-E-B. The TOPO reaction was trans-

formed into *E. coli* Top10 cells. Colonies were screened with M13 Forward and M13 Reverse primers by PCR amplification. Positive clones generated a 2.6 kbp product. Clones of pCR4Blunt-TOPO-E-B were submitted for sequencing with primers M13 Forward and Reverse, N62SeqF2, N62SeqF3, N62SeqF4, N63SeqR1, N63SeqR2, N63SeqR3, N11Seq F1 and N11Seq F2, N12SeqR1 and N12SeqR2.

Plasmid pCR4Blunt-TOPO-E-B was digested with HpaI and AatII to release a 2.4 kbp fragment. The E-B fragment was treated with Klenow polymerase to blunt the end and then was gel-purified. Plasmid pHT01-ald was digested with AatII and treated with Klenow polymerase to blunt the ends. The vector was then treated with calf intestinal alkaline phosphatase and was gel-purified. The E-B fragment was ligated to the linearized vector pHT01-ald, transformed into *E. coli* Top10 cells, and selected on LB plates containing 100 μg/mL ampicillin. Transformants were screened by PCR amplification with primers N3SeqF1 and N63SeqR1 to give a 2.4 kbp product. The resulting plasmid, pHT01-ald-EB, was transformed into JM103 cells, a *recA*⁺ *E. coli* strain. Plasmids prepared from *recA*⁺ strains form more multimers than *recA*⁻ strains. *Bacillus subtilis* transforms more efficiently with plasmid multimers rather than monomers (*Methods for General and Molecular Bacteriology*, supra). Plasmid DNA was prepared from JM103 and transformed into competent *B. subtilis* Δ*sacB*::T-H-C::erm #28 forming strain *B. subtilis* Δ*sacB*::T-H-C::erm #28/pHT01-ald-EB. Competent cells were prepared and transformed as previously described. Transformants were selected on LB plates containing 5 μg/mL chloramphenicol and screened by colony PCR with the primers N31 SeqF1 and N63SeqR4 for a 1.3 kbp product.

In the alternate cloning strategy, pCR4Blunt-TOPO-E-B was digested with HpaI and AatII releasing a 2.4 kbp fragment that was gel-purified. Plasmid pCR4-TOPO-B/A-ald was digested with HpaI and AatII and the 5.4 kbp vector fragment was gel-purified. The vector fragment from pCR4-TOPO-B/A-ald was ligated with the HpaI-AatII E-B fragment creating pCR4-TOPO-ald-EB. The ligation was transformed into *E. coli* Top10 cells and the resulting transformants were screened by PCR amplification with primers N11 SeqF2 and N63SeqR4 for a 2.1 kbp product. Plasmid pCR4-TOPO-ald-EB was digested with BamHI and AatII and SphI. The BamHI/AatII digest releases a 3.9 kbp ald-EB fragment that was gel-purified. The purpose of the SphI digest was to cut the remaining vector into smaller fragments so that it would not co-migrate on a gel with the ald-EB insert. Vector pHT01 was digested with BamHI and AatII and the 7.9 kbp vector fragment was gel-purified. The vector and ald-EB insert fragments were ligated to form plasmid pHT01-AEB and transformed into *E. coli* Top10 cells. Colonies were screened by PCR amplification with primers N62SeqF4 and HT R for a 1.5 kbp product. Plasmid was prepared and transformed into JM103. Plasmid DNA was prepared from JM103 and transformed into competent *B. subtilis* Δ*sacB*::T-H-C::erm #28 forming strain *B. subtilis* Δ*sacB*::T-H-C::erm #28/pHT01-AEB. Competent BE1010 cells were prepared and transformed as previously described. *Bacillus* transformants were screened by PCR amplification with primers N31 SeqF1 and N63SeqR4 for a 1.3 kbp product.

Demonstration of 1-butanol Production from Recombinant *B. subtilis*.

Three independent isolates of each strain of *B. subtilis* Δ*sacB*::T-H-C::erm #28/pHT01-ald-EB and *B. subtilis* Δ*sacB*::T-H-C::erm #28/pHT01-AEB were inoculated into shake flasks (approximately 175 mL total volume) containing 15 mL of medium. A *B. subtilis* BE1010 strain lacking the

exogenous 1-butanol, six gene pathway was also included as a negative control. The medium contained (per liter): 10 mL of 1 M (NH₄)₂SO₄; 5 mL of 1 M potassium phosphate buffer, pH 7.0; 100 mL of 1 M MOPS/KOH buffer, pH 7.0; 20 mL of 1 M L-glutamic acid, potassium salt; 10 g glucose; 10 mL of 5 g/L each of L-methionine, L-tryptophan, and L-lysine; 0.1 g each of yeast extract and casamino acids; 20 mL of metal mix; and appropriate antibiotics (5 mg chloramphenicol and erythromycin for the recombinant strains). The metal mix contained 200 mM MgCl₂, 70 mM CaCl₂, 5 mM MnCl₂, 0.1 mM FeCl₃, 0.1 mM ZnCl₂, 0.2 mM thiamine hydrochloride, 172 μM CuSO₄, 253 μM CoCl₂, and 242 μM Na₂MoO₄. The flasks were inoculated at a starting OD₆₀₀ of ≤0.1 units, sealed with non-vented caps, and incubated at 37° C. with shaking at approximately 200 rpm.

Approximately 24 h after inoculation, an aliquot of the broth was analyzed by HPLC (Shodex Sugar SH1011 column) with refractive index (RI) detection and GC (Varian CP-WAX 58(FFAP) CB column, 0.25 mm×0.2 μm×25 m) with flame ionization detection (FID) for 1-butanol content, as described in the General Methods section. The results of the 1-butanol determinations are given in Table 11.

TABLE 11

Production of 1-butanol by strains <i>B. subtilis</i> Δ <i>cscB</i> ::T-H-C::erm #28/pHT01-ald-EB and <i>B. subtilis</i> Δ <i>cscB</i> ::T-H-C::erm #28/pHT01-AEB		
Strain	1-butanol, HPLC RI peak area	1-butanol, mM*
BE1010 control	Not detected	Not detected
pHT01-ald-EB a	4629	0.19
pHT01-ald-EB b	3969	Not determined
pHT01-ald-EB c	4306	Not determined
pHT01-AEB a	4926	0.16
pHT01-AEB b	3984	Not determined
pHT01-AEB c	3970	Not determined

*Concentration determined by GC.

Strain suffixes "a", "b", and "c" indicate separate isolates.

Example 15

Production of 1-butanol from Glucose or Sucrose by Recombinant *E. coli*

To endow *E. coli* MG1655 with the ability to use sucrose as the carbon and energy source for 1-butanol production, a sucrose utilization gene cluster (*cscBKA*) from plasmid pScr1 (described below) was subcloned into pBHR-Ptrc-ald(opt) (described in Example 13) in this organism. The sucrose utilization genes (*cscA*, *cscK*, and *cscB*) encode a sucrose hydrolase (*CscA*), given as SEQ ID NO:157, D-fructokinase (*CscK*), given as SEQ ID NO:158, and sucrose permease (*CscB*), given as SEQ ID NO:159. To allow constitutive expression of the three genes from their natural promoter, the sucrose-specific repressor gene, *cscR*, that regulates the gene cluster is not present in the construct.

Cloning and expression of the sucrose utilization gene cluster *cscBKA* in plasmid pBHR-Ptrc-ald(opt)

The sucrose utilization gene cluster *cscBKA*, given as SEQ ID NO:156, was isolated from genomic DNA of a sucrose-utilizing *E. coli* strain derived from *E. coli* strain ATCC 13281. The genomic DNA was digested to completion with BamHI and EcoRI. Fragments having an average size of about 4 kbp were isolated from an agarose gel, ligated to plasmid pLitmus28 (New England Biolabs, Beverly, Mass.), which was then digested with BamHI and EcoRI. The result-

ing DNA was transformed into ultracompetent *E. coli* TOP10F' (Invitrogen, Carlsbad, Calif.). The transformants were plated on MacConkey agar plates containing 1% sucrose and 100 μg/mL ampicillin and screened for purple colonies. Plasmid DNA was isolated from the purple transformants and sequenced using primers M13 Forward (SEQ ID NO:45), M13 Reverse (SEQ ID NO:46), scr1 (SEQ ID NO:160), scr2 (SEQ ID NO:161), scr3 (SEQ ID NO:162), and scr4 (SEQ ID NO:163). The plasmid containing *cscB*, *cscK*, and *cscA* (*cscBKA*) genes was designated pScr1. Plasmid pScr1 was digested with XhoI and treated with the Klenow fragment of DNA polymerase to make blunt ends. The plasmid was then digested with AgeI, and the 4,179 bp *cscBKA* gene cluster fragment was gel-purified. Plasmid pBHR-Ptrc-ald(opt) was prepared as described in Example 13 and was digested with AgeI and NaeI. The resulting 6,003 bp pBHR-Ptrc-ald(opt) fragment was gel-purified. The *cscBKA* fragment was ligated with the pBHR-Ptrc-ald(opt), yielding pBHR-Ptrc-ald(opt)-*cscAKB*. Plasmid pBHR-Ptrc-ald(opt)-*cscAKB* was transformed into *E. coli* NovaXG electrocompetent cells (Novagen, Madison, Wis.) and sucrose utilization was confirmed by plating the transformants on McConkey agar plates containing 2% sucrose and 25 μg/mL kanamycin. In the pBHR-Ptrc-ald(opt)-*cscAKB* construct, the sucrose utilization genes were cloned downstream of Ptrc-ald(opt) as a separate fragment in the order *cscA*, *cscK*, and *cscB*.

Alternatively, the sucrose utilization genes were cloned in the opposite direction in pBHR-Ptrc-ald(opt). Plasmid pBHR-Ptrc-ald(opt) was digested with ScaI and AgeI, and the 5,971 bp pBHR-Ptrc-ald(opt) fragment was gel-purified. The 4,179 bp *cscBKA* fragment, prepared as described above, was ligated with the pBHR-Ptrc-ald(opt) fragment, yielding pBHR-Ptrc-ald(opt)-*cscBKA*. Plasmid pBHR-Ptrc-ald(opt)-*cscBKA* was transformed into *E. coli* NovaXG electrocompetent cells (Novagen, Madison, Wis.) and sucrose utilization was confirmed by plating the transformants on McConkey agar plates containing 2% sucrose and 25 μg/mL kanamycin. In the pBHR-Ptrc-ald(opt)-*cscBKA* construct, the sucrose utilization genes were cloned as a separate fragment downstream of Ptrc-ald(opt) in the order *cscB*, *cscK*, and *cscA*.

Demonstration of 1-butanol Production from Glucose or Sucrose Using Recombinant *E. coli*

E. coli strain MG1655 1.5GI-yqhD::Cm (described in Example 13) was transformed with plasmids pTrc99a-E-C-H-T (prepared as described in Example 13) and pBHR-Ptrc-ald(opt)-*cscAKB* or pBHR-Ptrc-ald(opt)-*cscBKA* to produce two strains, MG1655 1.5GI-yqhD::Cm/pTrc99a-E-C-H-T/pBHR-Ptrc-ald(opt)-*cscAKB* #9 and MG1655 1.5GI-yqhD::Cm/pTrc99a-E-C-H-T/pBHR-Ptrc-ald(opt)-*cscBKA* #1. Starter cultures of the two strains were prepared by growing the cells in LB medium containing 25 μg/mL of kanamycin and 100 μg/mL of carbenicillin. These cells were then used to inoculate shake flasks (approximately 175 mL total volume) containing 50, 70 and 150 mL of TM3a/glucose medium (with appropriate antibiotics) to represent high, medium and low oxygen conditions, respectively, as described in Example 13. A third strain, *E. coli* MG1655/pScr1, grown in TM3a/glucose medium containing 100 μg/mL carbenicillin, was used as a negative control. For each of the strains, an identical set of flasks was prepared with TM3a/sucrose medium (with appropriate antibiotics). TM3a/sucrose medium is identical to TM3a/glucose medium except that sucrose (10 g/L) replaces glucose. The flasks were inoculated at a starting OD₅₅₀ of ≤0.03 units and incubated as described in Example 13. With the exception of the negative control flasks, IPTG was added to the flasks (final concentra-

49

tion of 0.04 mM) when the cultures reached an OD₅₅₀ between 0.2 and 1.8 units. The cells were harvested when the OD₅₅₀ of the cultures increased at least 3-fold.

Approximately 24 h after inoculation, an aliquot of the broth was analyzed by HPLC (Shodex Sugar SH1011 column) with refractive index (RI) detection and GC(HP-INNOWax column, 30 m×0.53 mm id, 1 µm film thickness) with flame ionization detection (FID) for 1-butanol content, as described in the General Methods section.

The concentrations of 1-butanol in cultures following growth in the glucose and sucrose-containing media are given in Table 12 and Table 13, respectively. Both recombinant *E. coli* strains containing the 1-butanol biosynthetic pathway produced 1-butanol from glucose and sucrose under all oxygen conditions, while the negative control strain produced no detectable 1-butanol.

TABLE 12

Production of 1-butanol from glucose by recombinant <i>E. coli</i> strains MG1655 1.5G1-yqhD::Cm/pTrec99a-E-C-H-T/pBHR-Ptrc-ald(opt)-cscAKB #9 and MG1655 1.5G1-yqhD::Cm/pTrec99a-E-C-H-T/pBHR-Ptrc-ald(opt)-cscBKA #1			
Strain	O ₂ Level	1-butanol, mM	molar yield, %
cscBKA #1	high	0.01	0.03
cscBKA #1	medium	0.20	0.43
cscBKA #1	low	0.07	0.21
cscAKB #9	high	0.01	0.02
cscAKB #9	medium	0.17	0.35
cscAKB #9	low	0.04	0.12
pScr1	high	Not detected	Not detected
pScr1	medium	Not detected	Not detected
pScr1	low	Not detected	Not detected

TABLE 13

Production of 1-butanol from sucrose by recombinant <i>E. coli</i> strains.			
Strain	O ₂ Level	1-butanol, mM	molar yield, %
cscBKA #1	high	0.02	0.10
cscBKA #1	medium	0.02	0.11
cscBKA #1	low	0.01	0.09
cscAKB #9	high	0.03	0.11
cscAKB #9	medium	0.03	0.15
cscAKB #9	low	0.02	0.10
pScr1	high	Not detected	Not detected
pScr1	medium	Not detected	Not detected
pScr1	low	Not detected	Not detected

Example 16

Production of 1-butanol from Sucrose Using Recombinant *B. subtilis*

This example describes the production of 1-butanol from sucrose using recombinant *Bacillus subtilis*. Two independent isolates of *B. subtilis* strain ΔsacB::T-H-C::erm #28/pHT01-ald-EB (Example 14) were examined for 1-butanol production essentially as described in Example 14. The strains were inoculated into shake flasks (approximately 175 mL total volume) containing either 20 mL or 100 mL of medium to simulate high and low oxygen conditions, respectively. Medium A was exactly as described in Example 14, except that glucose was replaced with 5 g/L of sucrose. Medium B was identical to the TM3a/glucose medium described in Example 13, except that glucose was replaced with 10 g/L sucrose and the medium was supplemented with

50

(per L) 10 mL of a 5 g/L solution of each of L-methionine, L-tryptophan, and L-lysine. The flasks were inoculated at a starting OD₅₅₀ of ≤0.1 units, capped with vented caps, and incubated at 34° C. with shaking at 300 rpm.

Approximately 24 h after inoculation, an aliquot of the broth was analyzed by GC(HP-INNOWax column, 30 m×0.53 mm id, 1.0 µm film thickness) with FID detection for 1-butanol content, as described in the General Methods section. The results of the 1-butanol determinations are given in Table 14. The recombinant *Bacillus* strain containing the 1-butanol biosynthetic pathway produced detectable levels of 1-butanol under high and low oxygen conditions in both media.

TABLE 14

Production of 1-butanol from sucrose by <i>B. subtilis</i> strain ΔsacB::T-H-C::erm #28/pHT01-ald-EB			
Strain	Medium	O ₂ Level	1-BuOH, mM ^{1,2}
none	A	Not applicable	Not detected
pHT01-ald-EB a	A	high	+
pHT01-ald-EB b	A	high	+
pHT01-ald-EB a	A	low	0.01
pHT01-ald-EB b	A	low	0.01
none	B	Not applicable	Not detected
pHT01-ald-EB a	B	high	+
pHT01-ald-EB b	B	high	+
pHT01-ald-EB a	B	low	0.04
pHT01-ald-EB b	B	low	0.03

¹Concentration determined by GC.

²“+” indicates qualitative presence of 1-butanol.

Strain suffixes “a” and “b” indicate separate isolates.

Example 17

Production of 1-butanol from Glucose and Sucrose Using Recombinant *Saccharomyces cerevisiae*

This Example describes the production of 1-butanol in the yeast *Saccharomyces cerevisiae*. Of the six genes encoding enzymes catalyzing the steps in the 1-butanol biosynthetic pathway, five were cloned into three compatible yeast 2 micron (2µ) plasmids and co-expressed in *Saccharomyces cerevisiae*. The “upper pathway” is defined as the first three enzymatic steps, catalyzed by acetyl-CoA acetyltransferase (thlA, thiolase), 3-hydroxybutyryl-CoA dehydrogenase (hbd), and crotonase (crt). The lower pathway is defined as the fourth (butyl-CoA dehydrogenase, ter) and the fifth (butylaldehyde dehydrogenase, aid) enzymatic steps of the pathway. The last enzymatic step of the 1-butanol pathway is catalyzed by alcohol dehydrogenase, which may be encoded by endogenous yeast genes, e.g., adhI and adhII.

Expression of genes in yeast typically requires a promoter, followed by the gene of interest, and a transcriptional terminator. A number of constitutive yeast promoters were used in constructing expression cassettes for genes encoding the 1-butanol biosynthetic pathway, including FBA, GPD, and GPM promoters. Some inducible promoters, e.g. GAL1, GAL10, CUP1 were also used in intermediate plasmid construction, but not in the final demonstration strain. Several transcriptional terminators were used, including FBAt, GPDt, GPMt, ERG10t, and GAL1t. Genes encoding the 1-butanol biosynthetic pathway were first subcloned into a yeast plasmid flanked by a promoter and a terminator, which yielded expression cassettes for each gene. Expression cassettes were

optionally combined in a single vector by gap repair cloning, as described below. For example, the three gene cassettes encoding the upper pathway were subcloned into a yeast 2 μ plasmid. The ter and ald genes were each expressed individually in the 2 μ plasmids. Co-transformation of all three plasmids in a single yeast strain resulted in a functional 1-butanol biosynthetic pathway. Alternatively, several DNA fragments encoding promoters, genes, and terminators were directly combined in a single vector by gap repair cloning.

Methods for Constructing Plasmids and Strains in Yeast *Saccharomyces cerevisiae*.

Basic yeast molecular biology protocols including transformation, cell growth, gene expression, gap repair recombination, etc. are described in *Methods in Enzymology*, Volume 194, *Guide to Yeast Genetics and Molecular and Cell Biology* (Part A, 2004, Christine Guthrie and Gerald R. Fink (Eds.), Elsevier Academic Press, San Diego, Calif.).

The plasmids used in this Example were *E. coli*-*S. cerevisiae* shuttle vectors, pRS423, pRS424, pRS425, and pRS426 (American Type Culture Collection, Rockville, Md.), which contain an *E. coli* replication origin (e.g., pMB1), a yeast 2 μ origin of replication, and a marker for nutritional selection. The selection markers for these four vectors are His3 (vector pRS423), Trp1 (vector pRS424), Leu2 (vector pRS425) and Ura3 (vector pRS426). These vectors allow strain propagation in both *E. coli* and yeast strains. A yeast haploid strain BY4741 (MATa his3 Δ 1 leu2 Δ 0 met15 Δ 0 ura3 Δ 0) (Research Genetics, Huntsville, Ala., also available from ATCC 201388) and a diploid strain BY4743 (MATa/alpha his3 Δ 1/his3 Δ 1 leu2 Δ 0/leu2 Δ 0 lys2 Δ 0/LYS2 MET15/met15 Δ 0 ura3 Δ 0/ura3 Δ 0) (Research Genetics, Huntsville, Ala., also available from ATCC 201390) were used as hosts for gene cloning and expression. Construction of expression vectors for genes encoding 1-butanol biosynthetic pathway enzymes were performed by either standard molecular cloning techniques in *E. coli* or by the gap repair recombination method in yeast.

The gap repair cloning approach takes advantage of the highly efficient homologous recombination in yeast. Typically, a yeast vector DNA is digested (e.g., in its multiple cloning site) to create a "gap" in its sequence. A number of insert DNAs of interest are generated that contain a ≥ 21 bp sequence at both the 5' and the 3' ends that sequentially overlap with each other, and with the 5' and 3' terminus of the vector DNA. For example, to construct a yeast expression vector for "Gene X", a yeast promoter and a yeast terminator are selected for the expression cassette. The promoter and terminator are amplified from the yeast genomic DNA, and Gene X is either PCR amplified from its source organism or obtained from a cloning vector comprising Gene X sequence. There is at least a 21 bp overlapping sequence between the 5' end of the linearized vector and the promoter sequence, between the promoter and Gene X, between Gene X and the terminator sequence, and between the terminator and the 3' end of the linearized vector. The "gapped" vector and the insert DNAs are then co-transformed into a yeast strain and plated on the SD minimal dropout medium, and colonies are selected for growth of cultures and mini preps for plasmid DNAs. The presence of correct insert combinations can be confirmed by PCR mapping. The plasmid DNA isolated from yeast (usually low in concentration) can then be transformed into an *E. coli* strain, e.g. TOP10, followed by mini preps and restriction mapping to further verify the plasmid construct. Finally the construct can be verified by sequence analysis. Yeast transformants of positive plasmids are grown in SD medium for performing enzyme assays to characterize the activities of the enzymes expressed by the genes of interest.

Yeast cultures were grown in YPD complex medium or Synthetic Minimal dropout medium containing glucose (SD medium) and the appropriate compound mixtures that allow complementation of the nutritional selection markers on the plasmids (*Methods in Enzymology*, Volume 194, *Guide to Yeast Genetics and Molecular and Cell Biology*, 2004, Part A, pp. 13-15). The sugar component in the SD drop out medium was 2% glucose. For 1-butanol production, yeast cultures were also grown in Synthetic Minimal dropout medium with 2% sucrose (SS medium).

For enzyme activity analysis, a single colony of each strain was streaked onto a fresh plate containing SD minimal drop out medium and incubated at 30° C. for 2 days. The cells on this plate were used to inoculate 20 mL of SD drop out medium and in a 125 mL shake flask and grown overnight at 30° C., with shaking at 250 rpm. The optical density (OD₆₀₀) of the overnight culture was measured, and the culture was diluted to an OD₆₀₀=0.1 in 250 mL of the same medium in a 1.0L shake flask, and grown at 30° C. with shaking at 250 rpm to an OD₆₀₀ of between 0.8 to 1.0. The cells were then harvested by centrifugation at 2000xg for 10 min, and resuspended in 20 mL of 50 mM Tris-HCl buffer, pH 8.5. Enzyme assays were performed as described above.

Construction of plasmid pNY102 for thlA and hbd co-expression.

A number of dual expression vectors were constructed for the co-expression of thlA and hbd genes. The *Saccharomyces cerevisiae* ERG10 gene is a functional ortholog of the thlA gene. Initially, a dual vector of ERG10 and hbd was constructed using the yeast GAL1-GAL10 divergent dual promoter, the GAL1 terminator (GAL1t) and the ERG10 terminator (ERG10t). The ERG10 gene-ERG10t DNA fragment was PCR amplified from genomic DNA of *Saccharomyces cerevisiae* strain BY4743, using primers OT731 (SEQ ID NO:164) and OT732 (SEQ ID NO:165). The yeast GAL1-GAL10 divergent promoter was also amplified by PCR from BY4743 genomic DNA using primers OT733 (SEQ ID NO:166) and OT734 (SEQ ID NO:167). The hbd gene was amplified from *E. coli* plasmid pTrc99a-E-C-H-T (described in Example 13) using PCR primers OT735 (SEQ ID NO:168) and OT736 (SEQ ID NO:169). GAL1t was amplified from BY4743 genomic DNA using primers OT737 (SEQ ID NO:170) and OT738 (SEQ ID NO:171). Four PCR fragments, erg10-ERG10t, GAL1-GAL10 promoters, hbd, and GAL1t, were thus obtained with approximately 25 bp overlapping sequences between each adjacent PCR fragment. GAL1t and ERG1E-ERG10t fragments each contain approximately 25 bp overlapping sequences with the yeast vector pRS425. To assemble these sequences by gap repair recombination, the DNA fragments were co-transformed into the yeast strain BY4741 together with vector pRS425 which was digested with BamHI and HindIII enzymes. Colonies were selected from SD-Leu minimal plates, and clones with inserts were identified by PCR amplification. The new plasmid was named pNY6 (pRS425.ERG 10t-erg10-GAL10-GAL1-hbd-GAL1 t). Further confirmation was performed by restriction mapping.

The yeast strain BY4741 (pNY6), prepared by transforming plasmid pNY6 into *S. cerevisiae* BY4741, showed good Hbd activity but no thiolase activity. Due to the lack of thiolase activity, the ERG10 gene was replaced with the thlA gene by gap repair recombination. The thlA gene was amplified from *E. coli* vector pTrc99a-E-C-H-T by PCR using primers OT797 (SEQ ID NO:172) which adds a SphI restriction site, and OT798 (SEQ ID NO:173) which adds an AscI restriction site. Plasmid pNY6 was digested with SphI and PstI restriction enzymes, gel-purified, and co-transformed into yeast

BY4741 along with the PCR product of *thlA*. Due to the 30 bp overlapping sequences between the PCR product of *thlA* and the digested pNY6, the *thlA* gene was recombined into pNY6 between the GAL10 promoter and the ERG10t terminator. This yielded plasmid pNY7 (pRS425.ERG10t-*thlA*-GAL10-GAL1-hbd-GAL1t), which was verified by PCR and restriction mapping.

In a subsequent cloning step based on gap repair recombination, the GAL10 promoter in pNY7 was replaced with the CUP1 promoter, and the GAL1 promoter was replaced with the strong GPD promoter. This plasmid, pNY10 (pRS425.ERG10t-*thlA*-CUP1-GPD-hbd-GAL1t) allows for the expression of the *thlA* gene under CUP1, a copper inducible promoter, and the expression of the *hbd* gene under the GPD promoter. The CUP1 promoter sequence was PCR amplified from yeast BY4743 genomic DNA using primers OT806 (SEQ ID NO:174), and OT807 (SEQ ID NO:175). The GPD promoter was amplified from BY4743 genomic DNA using primers OT808 (SEQ ID NO:176) and OT809 (SEQ ID NO:177). PCR products of the CUP1 and the GPD promoters were combined with pNY7 plasmid digested with NcoI and SphI restriction enzymes. From this gap repair cloning step, plasmid pNY10 was constructed, which was verified by PCR and restriction mapping. Yeast BY4741 strain containing pNY10 had Hbd activity, but no ThlA activity. The Hbd activity under GPD promoter was significantly improved compared to the GALL promoter controlled Hbd activity (1.8 U/mg vs. 0.40 U/mg). Sequencing analysis revealed that the *thlA* gene in pNY10 had a one base deletion near the 3' end, which resulted in a truncated protein. This explains the lack of thiolase activity in the strain.

Plasmid pNY12 was constructed with the correct *thlA* gene sequence. The *thlA* gene was cut from the vector pTrc99a-E-C-H-T by digestion with SphI and AseI. The FBA1 promoter was PCR amplified from BY4743 genomic DNA using primers OT799 (SEQ ID NO:178) and OT761 (SEQ ID NO:179), and digested with SalI and SphI restriction enzymes. The *thlA* gene fragment and FBA1 promoter fragment were ligated into plasmid pNY10 at AseI and SalI sites, generating plasmid pNY12 (pRS425.ERG10t-*thlA*-FBA1), which was confirmed by restriction mapping. pNY12 was transformed into yeast strain BY4741 and the resulting transformant showed a ThlA activity of 1.66 U/mg.

The FBA1 promoter-*thlA* gene fragment from pNY12 was re-subcloned into pNY10. The pNY10 vector was cut with the AseI restriction enzyme and ligated with the AseI digested FBA1 promoter-*thlA* gene fragment isolated from plasmid pNY12. This created a new plasmid with two possible insert orientations. The clones with FBA1 and GPD promoters located adjacent to each other in opposite orientation were chosen and this plasmid was named pNY102. pNY102 (pRS425.ERG10t-*thlA*-FBA1-GPD-hbd-GAL1t) was verified by restriction mapping. Strain DPD5206 was made by transforming pNY102 into yeast strain BY4741. The ThlA activity of DPD5206 was 1.24 U/mg and the Hbd activity was 0.76 U/mg.

Construction of Plasmid pNY11 for Crt Expression.

The crt gene expression cassette was constructed by combining the GPM1 promoter, the crt gene, and the GPM1t terminator into vector pRS426 using gap repair recombination in yeast. The GPM1 promoter was PCR amplified from yeast BY4743 genomic DNA using primers OT803 (SEQ ID NO:180) and OT804 (SEQ ID NO:181). The crt gene was amplified using PCR primers OT785 (SEQ ID NO:182) and OT786 (SEQ ID NO:183) from *E. coli* plasmid pTrc99a-E-C-H-T. The GPM1t terminator was PCR amplified from yeast BY4743 genomic DNA using OT787 (SEQ ID NO:184) and

OT805 (SEQ ID NO:185). Yeast vector pRS426 was digested with BamHI and HindIII and was gel-purified. This DNA was co-transformed with the PCR products of the GPM1 promoter, the crt gene and the GPM1 terminator into yeast BY4741 competent cells. Clones with the correct inserts were verified by PCR and restriction mapping and the resulting yeast strain BY4741 (pNY11: pRS426-GPM1-crt-GPM1t) had a Crt activity of 85 U/mg.

Construction of plasmid pNY103 for *thlA*, *hbd* and *cdt* co-expression. For the co-expression of the upper 1-butanol pathway enzymes, the crt gene cassette from pNY11 was subcloned into plasmid pNY102 to create an *hbd*, *thlA*, and crt expression vector. A 2,347 bp DNA fragment containing the GPM1 promoter, the crt gene, and the GPM1 terminator was cut from plasmid pNY11 with SacI and NotI restriction enzymes and cloned into vector pNY102, which was digested with NotI and partially digested with SacI, producing the expression vector pNY103 (pRS425.ERG10t-*thlA*-FBA1-GPD-hbd-GAL1t-GPM1t-crt-GPM1t). Following confirmation of the presence of all three cassettes in pNY103 by digestion with HindIII, the plasmid was transformed into yeast BY4743 cells and the transformed yeast strain was named DPD5200. When grown under standard conditions, DPD5200 showed ThlA, Hbd, and Crt enzyme activities of 0.49 U/mg, 0.21 U/mg and 23.0 U/mg, respectively.

Construction of Plasmid pNY8 for Ald Expression.

A codon optimized gene named tery (SEQ ID NO:186), encoding the Ter protein (SEQ ID NO:187), and a codon optimized gene named aldy (SEQ ID NO:188), encoding the Ald protein (SEQ ID NO:189) were synthesized using preferred codons of *Saccharomyces cerevisiae*. Plasmid pTERy containing the codon optimized ter gene and pALDy containing the codon optimized ald gene were made by DNA2.0 (Palo Alto, Calif.).

To assemble pNY8 (pRS426.GPD-ald-GPDt), three insert fragments including a PCR product of the GPD promoter (synthesized from primers OT800 (SEQ ID NO:190) and OT758, (SEQ ID NO:191), and BY4743 genomic DNA), an aldy gene fragment excised from pALDy by digestion with NcoI and SfiI (SEQ ID NO:188), and a PCR product of the GPD terminator (synthesized from primers OT754 (SEQ ID NO:192) and OT755 (SEQ ID NO:193), and BY4743 genomic DNA) were recombined with the BamHI, HindIII digested pRS426 vector via gap repair recombination cloning. Yeast BY4741 transformation clones were analyzed by PCR mapping. The new plasmid thus constructed, pNY8, was further confirmed by restriction mapping. The yeast BY4741 transformants containing pNY8 were analyzed for Ald activity and the specific activity towards butyryl-CoA was approximately 0.07 U/mg.

Construction of Plasmids pNY9 and pNY13 for Ter Expression.

The codon optimized tery gene was cloned into vector pRS426 under control of the FBA1 promoter by gap repair cloning. The FBA1 promoter was PCR amplified from yeast BY4743 genomic DNA using primers OT760 (SEQ ID NO:194) and OT792 (SEQ ID NO:195). The tery gene was obtained by digestion of plasmid pTERy by SphI and NotI restriction enzymes that resulted in the fragment given as SEQ ID NO:186. The PCR fragment of FBA1 terminator was generated by PCR from yeast BY4743 genomic DNA using primers OT791 (SEQ ID NO:196) and OT765 (SEQ ID NO:197). Three DNA fragments, the FBA1 promoter, the tery gene and the FBA1 terminator, were combined with the BamHI, HindIII digested pRS426 vector and transformed into yeast BY4741 by gap repair recombination. The resulting plasmid, pNY9 (pRS426-FBA1-tery-FBA1t) was con-

55

firmed by PCR mapping, as well as restriction digestion. The yeast BY4741 transformant of pNY9 produced a Ter activity of 0.26 U/mg.

To make the final 1-butanol biosynthetic pathway strain, it was necessary to construct a yeast expression strain that contained several plasmids, each with a unique nutritional selection marker. Since the parent vector pRS426 contained a Ura selection marker, the ter expression cassette was subcloned into vector pRS423, which contained a His3 marker. A 3.2 kb fragment containing the FBA1-tery-FBA1t cassette was isolated from plasmid pNY9 by digestion with SacI and XhoI restriction enzymes, and ligated into vector pRS423 that was cut with these same two enzymes. The new plasmid, pNY13 (pRS423—FBA1-tery-FBA1t) was mapped by restriction digestion. pNY13 was transformed into BY4741 strain and the transformant was cultured in SD-His medium, yielding a strain with a Ter activity of 0.19 U/mg.

Construction of a Yeast Strain Containing 1-butanol Biosynthetic Pathway Genes for Demonstration of 1-butanol Production.

As described above, yeast strain DPD5200 was constructed by transformation of plasmid pNY103 into *S. cerevisiae* strain BY4743, which allows co-expression of thlA, hbd and crt genes. Yeast competent cells of DPD5200 were prepared as described above, and plasmids pNY8 and pNY13 were co-transformed into DPD5200, generating strain DPD5213. DPD5213 allows for the simultaneous constitutive expression of five genes in the 1-butanol biosynthetic pathway, thlA, hbd, crt, ter and aid. Strain DPD5212 (*S. cerevisiae* strain BY4743 transformed with empty plasmids, pRS425 and pRS426) was used as a negative control. Four independent isolates of strain DPD5213 were grown on SD-Ura-Leu-His dropout minimal medium in the presence of either 2% glucose or 2% sucrose to allow the growth complementation of all three plasmids. A single isolate of DPD5212 was similarly grown in appropriate medium.

To demonstrate 1-butanol production by aerobic cultures, a single colony of each strain was streaked onto a fresh agar plate containing SD minimal drop out growth medium (containing 2% glucose) or SS minimal drop out growth medium (containing 2% sucrose) and incubated at 30° C. for 2 days. Cells from these plates were used to inoculate 20 mL of the minimal drop out medium (either SD or SS) in 125 mL plastic shake flasks and were grown overnight at 30° C. with shaking at 250 rpm. The optical density (OD₆₀₀) of the overnight culture was measured, the culture was diluted to OD₆₀₀ of 0.1 in 25 mL of the same medium in a 125 mL shake flask, and grown at 30° C. with shaking at 250 rpm.

Aliquots of the culture were removed at 24 h and 48 h for GC analysis of 1-butanol production (HP-INNOWax column, 30 m×0.53 mm id, 1 µm film thickness) with FID detection, as described in the General Methods section. The results of the GC analysis are given in Table 15.

TABLE 15

Production of 1-butanol from glucose and sucrose by <i>S. cerevisiae</i> strain DPD5213			
Strain ¹	Sugar	1-butanol at 24 h, mg/L ²	1-butanol at 48 h, mg/L ²
DPD5212	Glucose	Not detected	Not detected
DPD5213 a	Glucose	0.4	0.5
DPD5213 b	Glucose	0.9	0.2
DPD5213 c	Glucose	1.0	0.6
DPD5213 d	Glucose	0.8	0.3
DPD5212	Sucrose	Not detected	Not detected

56

TABLE 15-continued

Production of 1-butanol from glucose and sucrose by <i>S. cerevisiae</i> strain DPD5213			
Strain ¹	Sugar	1-butanol at 24 h, mg/L ²	1-butanol at 48 h, mg/L ²
DPD5213 a	Sucrose	Not detected	1.7
DPD5213 b	Sucrose	Not detected	1.3
DPD5213 c	Sucrose	0.2	1.5
DPD5213 d	Sucrose	0.6	0.9

¹Independent isolates are indicated by a-d.

²Concentration determined by GC.

Example 18

Prophetic

Expression of the 1-butanol Biosynthetic Pathway in
Lactobacillus plantarum

The purpose of this prophetic Example is to describe how to express the 1-butanol biosynthetic pathway in *Lactobacillus plantarum*. The six genes of the 1-butanol pathway, encoding six enzyme activities, are divided into two operons for expression. The first three genes of the pathway (thl, hbd, and crt, encoding the enzymes acetyl-CoA acetyltransferase, 3-hydroxybutyryl-CoA dehydrogenase, and crotonase, respectively) are integrated into the chromosome of *Lactobacillus plantarum* by homologous recombination using the method described by Hols et al. (*Appl. Environ. Microbiol.* 60:1401-1413 (1994)). The last three genes (EgTER, aid, and bdhB, encoding the enzymes butyryl-CoA dehydrogenase, butyraldehyde dehydrogenase and butanol dehydrogenase, respectively) are cloned into an expression plasmid and transformed into the *Lactobacillus* strain carrying the integrated upper pathway 1-butanol genes. *Lactobacillus* is grown in MRS medium (Difco Laboratories, Detroit, Mich.) at 37° C. Chromosomal DNA is isolated from *Lactobacillus plantarum* as described by Moreira et al. (*BMC Microbiol.* 5:15 (2005)).

Integration.

The thl-hbd-crt cassette under the control of the synthetic P11 promoter (Rud et al., *Microbiology* 152:1011-1019 (2006)) is integrated into the chromosome of *Lactobacillus plantarum* ATCC BAA-793 (NCIMB 8826) at the IdhL1 locus by homologous recombination. To build the IdhL integration targeting vector, a DNA fragment from *Lactobacillus plantarum* (Genbank NC_004567) with homology to IdhL is PCR amplified with primers LDH EcoRV F (SEQ ID NO:198) and LDH AatIIR (SEQ ID NO:199). The 1986 bp PCR fragment is cloned into pCR4Blunt-TOPO and sequenced. The pCR4Blunt-TOPO-IdhL1 clone is digested with EcoRV and AatII releasing a 1982 bp IdhL1 fragment that is gel-purified. The integration vector pFP988, described in Example 14, is digested with HindIII and treated with Klenow DNA polymerase to blunt the ends. The linearized plasmid is then digested with AatII and the 2931 bp vector fragment is gel-purified. The EcoRV/AatII IdhL1 fragment is ligated with the pFP988 vector fragment and transformed into *E. coli* Top10 cells. Transformants are selected on LB agar plates containing ampicillin (100 µg/mL) and are screened by colony PCR to confirm construction of pFP988-IdhL.

To add a selectable marker to the integrating DNA, the Cm gene with its promoter is PCR amplified from pC194 (Genbank NC_002013) with primers Cm F (SEQ ID NO:200) and Cm R (SEQ ID NO:201), amplifying a 836 bp PCR product.

The amplicon is cloned into pCR4Blunt-TOPO and transformed into *E. coli* Top10 cells, creating pCR4Blunt-TOPO-Cm. After sequencing to confirm that no errors are introduced by PCR, the Cm cassette is digested from pCR4Blunt-TOPO-Cm as an 828 bp MluI/SwaI fragment and is gel-purified. The IdhL-homology containing integration vector pFP988-IdhL is digested with MluI and SwaI and the 4740 bp vector fragment is gel-purified. The Cm cassette fragment is ligated with the pFP988-IdhL vector creating pFP988-DldhL::Cm.

Finally the thl-hbd-crt cassette from pFP988Dss-T-H-C, described in Example 14, is modified to replace the amylase promoter with the synthetic P11 promoter. Then, the whole operon is moved into pFP988-DldhL::Cm. The P11 promoter is built by oligonucleotide annealing with primer P11 F (SEQ ID NO:202) and P11 R (SEQ ID NO:203). The annealed oligonucleotide is gel-purified on a 6% Ultra PAGE gel (Embi Tec, San Diego, Calif.). The plasmid pFP988Dss-T-H-C is digested with XhoI and SmaI and the 9 kbp vector fragment is gel-purified. The isolated P11 fragment is ligated with the digested pFP988Dss-T-H-C to create pFP988-P11-T-H-C. Plasmid pFP988-P11-T-H-C is digested with XhoI and BamHI and the 3034 bp P11-T-H-C fragment is gel-purified. pFP988-DldhL::Cm is digested with XhoI and BamHI and the 5558 bp vector fragment isolated. The upper pathway operon is ligated with the integration vector to create pFP988-DldhL-P11-THC::Cm.

Integration of pFP988-DldhL-P11-THC::Cm into *L. plantarum* BAA-793 to Form *L. plantarum* ΔldhL1::T-H-C::Cm Comprising Exogenous thl, hbd, and crt Genes.

Electrocompetent cells of *L. plantarum* are prepared as described by Aukrust, T. W., et al. (In: *Electroporation Protocols for Microorganisms*; Nickoloff, J. A., Ed.; *Methods in Molecular Biology*, Vol. 47; Humana Press, Inc., Totowa, N.J., 1995, pp 201-208). After electroporation, cells are outgrown in MRSSM medium (MRS medium supplemented with 0.5 M sucrose and 0.1 M MgCl₂) as described by Aukrust et al. supra for 2 h at 37° C. without shaking. Electroporated cells are plated for selection on MRS plates containing chloramphenicol (10 μg/mL) and incubated at 37° C. Transformants are initially screened by colony PCR amplification to confirm integration, and initial positive clones are then more rigorously screened by PCR amplification with a battery of primers.

Plasmid Expression of EgTER, Ald, and bdhB Genes.

The three remaining 1-butanol genes are expressed from plasmid pTRKH3 (O'Sullivan D J and Klaenhammer T R, *Gene* 137:227-231 (1993)) under the control of the *L. plantarum* IdhL promoter (Ferain et al., *J. Bacteriol.* 176:596-601 (1994)). The IdhL promoter is PCR amplified from the genome of *L. plantarum* ATCC BAA-793 with primers P IdhL F (SEQ ID NO:204) and P IdhL R (SEQ ID NO:205). The 369 bp PCR product is cloned into pCR4Blunt-TOPO and sequenced. The resulting plasmid, pCR4Blunt-TOPO-PldhL is digested with SacI and BamHI releasing the 359 bp PldhL fragment.

pHT01-ald-EB, described in Example 14, is digested with SacI and BamHI and the 10503 bp vector fragment is recovered by gel purification. The PldhL fragment and vector are ligated creating pHT01-PldhL-ald-EB.

To subclone the IdhL promoter-ald-EgTER-bdh cassette, pHT01-PldhL-ald-EB is digested with MluI and the ends are treated with Klenow DNA polymerase. The linearized vector is digested with SalI and the 4270 bp fragment containing the PldhL-AEB fragment is gel-purified. Plasmid pTRKH3 is digested with SalI and EcoRV and the gel-purified vector fragment is ligated with the PldhL-AEB fragment. The ligation mixture is transformed into *E. coli* Top 10 cells and

transformants are plated on Brain Heart Infusion (BHI, Difco Laboratories, Detroit, Mich.) plates containing erythromycin (150 mg/L). Transformants are screened by PCR to confirm construction of pTRKH3-ald-E-B. The expression plasmid, pTRKH3-ald-E-B is transformed into *L. plantarum* ΔldhL1::T-H-C::Cm by electroporation, as described above.

L. plantarum ΔldhL1::T-H-C::Cm containing pTRKH3-ald-E-B is inoculated into a 250 mL shake flask containing 50 mL of MRS medium plus erythromycin (10 μg/mL) and grown at 37° C. for 18 to 24 h without shaking. After 18 h to 24, 1-butanol is detected by HPLC or GC analysis, as described in the General Methods section.

Example 19

Prophetic

Expression of the 1-butanol Biosynthetic Pathway in *Enterococcus faecalis*

The purpose of this prophetic Example is to describe how to express the 1-butanol biosynthetic pathway in *Enterococcus faecalis*. The complete genome sequence of *Enterococcus faecalis* strain V583, which is used as the host strain for the expression of the 1-butanol biosynthetic pathway in this Example, has been published (Paulsen et al., *Science* 299: 2071-2074 (2003)). Plasmid pTRKH3 (O'Sullivan D J and Klaenhammer T R, *Gene* 137:227-231 (1993)), an *E. coli*/Gram-positive shuttle vector, is used for expression of the six genes (thlA, hbd, crt, EgTER, aid, bdhB) of the 1-butanol pathway in one operon. pTRKH3 contains an *E. coli* plasmid p15A replication origin and the pAMP1 replicon, and two antibiotic resistance selection markers, tetracycline resistance and erythromycin resistance. Tetracycline resistance is only expressed in *E. coli*, and erythromycin resistance is expressed in both *E. coli* and Gram-positive bacteria. Plasmid pAMP1 derivatives can replicate in *E. faecalis* (Poyart et al., *FEMS Microbiol. Lett.* 156:193-198 (1997)). The inducible nisA promoter (PnisA), which has been used for efficient control of gene expression by nisin in a variety of Gram-positive bacteria including *Enterococcus faecalis* (Eichenbaum et al., *Appl. Environ. Microbiol.* 64:2763-2769 (1998)), is used to control expression of the six desired genes encoding the enzymes of the 1-butanol biosynthetic pathway.

The linear DNA fragment (215 bp) containing the nisA promoter (Chandrapati et al., *Mol. Microbiol.* 46(2):467-477 (2002)) is PCR-amplified from *Lactococcus lactis* genomic DNA with primers F-PnisA(EcoRV) (SEQ ID NO:206) and R-PnisA(PmeI BamHI) (SEQ ID NO:207). The 215 bp PCR fragment is digested with EcoRV and BamHI, and the resulting PnisA fragment is gel-purified. Plasmid pTRKH3 is digested with EcoRV and BamHI and the vector fragment is gel-purified. The linearised pTRKH3 is ligated with the PnisA fragment. The ligation mixture is transformed into *E. coli* Top10 cells by electroporation and transformants are selected following overnight growth at 37° C. on LB agar plates containing erythromycin (25 μg/mL). The transformants are then screened by colony PCR with primers F-PnisA (EcoRV) and R-PnisA(BamHI) to confirm the correct clone of pTRKH3-PnisA.

Plasmid pTRKH3-PnisA is digested with PmeI and BamHI, and the vector is gel-purified. Plasmid pHT01-ald-EgTER-bdhB is constructed as described in example 14 and is digested with SmaI and BamHI, and the 2,973 bp ald-EgTER-bdhB fragment is gel-purified. The 2,973 bp ald-EgTER-bdhB fragment is ligated into the pTRKH3-PnisA vector at the PmeI and BamHI sites. The ligation mixture is

transformed into *E. coli* Top10 cells by electroporation and transformants are selected following incubation at 37° C. overnight on LB agar plates containing erythromycin (25 µg/mL). The transformants are then screened by colony PCR with primers ald forward primer N27F1 (SEQ ID NO: 31) and bdhB reverse primer N65 (SEQ ID NO: 44). The resulting plasmid is named pTRKH3-PnisA-ald-EgTER-bdhB (=pTRKH3-A-E-B).

Plasmid pTRKH3-A-E-B is purified from the transformant and used for further cloning of the remaining genes (thlA, hbd, crt) into the BamHI site located downstream of the bdhB gene. Plasmid pTRKH3-A-E-B is digested with BamHI and treated with the Klenow fragment of DNA polymerase to make blunt ends. Plasmid pFP988Dss-thlA-hbd-crt (=pFP988Dss-T-H—C) is constructed as described in Example 14 and is digested with SmaI and BamHI. The resulting 2,973 bp thlA-hbd-crt fragment is treated with the Klenow fragment of DNA polymerase to make blunt ends and is gel-purified. The 2,973 bp thlA-hbd-crt fragment is ligated with the linearised pTRKH3-A-E-B. The ligation mixture is transformed into *E. coli* Top10 cells by electroporation and transformants are selected following overnight growth at 37° C. on LB agar plates containing erythromycin (251 µg/mL). The transformants are then screened by colony PCR with primers thlA forward primer N7 (SEQ ID NO: 21) and crt reverse primer N4 (SEQ ID NO: 18). The resulting plasmid is named pTRKH3-PnisA-ald-EgTER-bdhB-thlA-hbd-crt

(=pTRKH3-A-E-B-T-H-C). Plasmid pTRKH3-A-E-B-T-H-C is prepared from the *E. coli* transformants and transformed into electro-competent *E. faecalis* V583 cells by electroporation using methods known in the art (Aukrust, T. W., et al. In: *Electroporation Protocols for Microorganisms*; Nickoloff, J. A., Ed.; *Methods in Molecular Biology*, Vol. 47; Humana Press, Inc., Totowa, N.J., 1995, pp 217-226), resulting in *E. faecalis* V583/pTRKH3-A-E-B-T-H-C.

The second plasmid containing nisA regulatory genes, nisR and nisK, the add9 spectinomycin resistance gene, and the pSH71 origin of replication is transformed into *E. faecalis* V583/pTRKH3-A-E-B-T-H-C by electroporation. The plasmid containing pSH71 origin of replication is compatible with pAMP1 derivatives in *E. faecalis* (Eichenbaum et al., supra). Double drug resistant transformants are selected on LB agar plates containing erythromycin (25 µg/mL) and spectinomycin (100 µg/mL).

The resulting *E. faecalis* strain V583B harboring two plasmids, i.e., an expression plasmid (pTRKH3-A-E-B-T-H-C) and a regulatory plasmid (pSH71-nisRK), is inoculated into a 250 mL shake flask containing 50 mL of Todd-Hewitt broth supplemented with yeast extract (0.2%) (Fischetti et al., *J. Exp. Med.* 161:1384-1401 (1985)), nisin (20 µg/mL) (Eichenbaum et al., supra), erythromycin (25 µg/mL), and spectinomycin (100 µg/mL). The flask is incubated without shaking at 37° C. for 18 to 24 h, after which time, 1-butanol production is measured by HPLC or GC analysis, as described in the General Methods section.

SEQUENCE LISTING

<160> NUMBER OF SEQ ID NOS: 208

<210> SEQ ID NO 1

<211> LENGTH: 1179

<212> TYPE: DNA

<213> ORGANISM: Clostridium acetobutylicum

<400> SEQUENCE: 1

```

atgaaagaag ttgtaatagc tagtgcagta agaacagcga ttggatctta tggaaagtct      60
cttaaggatg taccagcagt agatttagga gctacagcta taaaggaagc agttaaaaaa     120
gcaggaataa aaccagagga tgtaaatgaa gtcattttag gaaatgttct tcaagcaggt     180
ttaggacaga atccagcaag acaggcatct tttaaagcag gattaccagt tgaaattcca     240
gctatgacta ttaataaggt ttgtggttca ggacttagaa cagttagctt agcagcacia     300
attataaaag caggagatgc tgacgtaata atagcagggtg gtatggaaaa tatgtctaga     360
gctccttact tagcgaataa cgctagatgg ggatatagaa tgggaaacgc taaatttgtt     420
gatgaaatga tcaactgacgg attgtgggat gcatttaatg attaccacat gggaataaca     480
gcagaaaaa tagctgagag atggaacatt tcaagagaag aacaagatga gtttgctctt     540
gcatcacaaa aaaaagctga agaagctata aaatcagggtc aatttaaaga tgaaatagtt     600
cctgtagtaa ttaaaggcag aaaggagaaa actgtagttg atacagatga gcaccctaga     660
tttgatcaa ctatagaagg acttgcaaaa taaaacctg cttcaaaaa agatggaaaca     720
gttacagctg gtaatgcac aggattaaat gactgtgcag cagtacttgt aatcatgagt     780
gcagaaaaa gctaaagagct tggagtaaaa ccacttgcta agatagtctt ttatgggttca     840
gcaggagtgt accagcaat aatgggatat ggacctttct atgcaacaaa agcagctatt     900
gaaaaagcag gttggacagt tgatgaatta gatttaatat aatcaaatga agcttttgca     960
gctcaaaagt tagcagtagc aaaagattta aaatttgata tgaataaagt aaatgtaaat    1020

```

-continued

```

ggaggagcta ttgcccttgg tcaccaatt ggagcatcag gtgcaagaat actcggtact 1080
cttgtacacg caatgcaaaa aagagatgca aaaaaaggct tagcaacttt atgtataggt 1140
ggcggacaag gaacagcaat attgctagaa aagtgctag 1179

```

```

<210> SEQ ID NO 2
<211> LENGTH: 392
<212> TYPE: PRT
<213> ORGANISM: Clostridium acetobutylicum

```

```

<400> SEQUENCE: 2

```

```

Met Lys Glu Val Val Ile Ala Ser Ala Val Arg Thr Ala Ile Gly Ser
 1             5             10             15
Tyr Gly Lys Ser Leu Lys Asp Val Pro Ala Val Asp Leu Gly Ala Thr
      20             25             30
Ala Ile Lys Glu Ala Val Lys Lys Ala Gly Ile Lys Pro Glu Asp Val
      35             40             45
Asn Glu Val Ile Leu Gly Asn Val Leu Gln Ala Gly Leu Gly Gln Asn
      50             55             60
Pro Ala Arg Gln Ala Ser Phe Lys Ala Gly Leu Pro Val Glu Ile Pro
      65             70             75             80
Ala Met Thr Ile Asn Lys Val Cys Gly Ser Gly Leu Arg Thr Val Ser
      85             90             95
Leu Ala Ala Gln Ile Ile Lys Ala Gly Asp Ala Asp Val Ile Ile Ala
      100            105            110
Gly Gly Met Glu Asn Met Ser Arg Ala Pro Tyr Leu Ala Asn Asn Ala
      115            120            125
Arg Trp Gly Tyr Arg Met Gly Asn Ala Lys Phe Val Asp Glu Met Ile
      130            135            140
Thr Asp Gly Leu Trp Asp Ala Phe Asn Asp Tyr His Met Gly Ile Thr
      145            150            155            160
Ala Glu Asn Ile Ala Glu Arg Trp Asn Ile Ser Arg Glu Glu Gln Asp
      165            170            175
Glu Phe Ala Leu Ala Ser Gln Lys Lys Ala Glu Glu Ala Ile Lys Ser
      180            185            190
Gly Gln Phe Lys Asp Glu Ile Val Pro Val Val Ile Lys Gly Arg Lys
      195            200            205
Gly Glu Thr Val Val Asp Thr Asp Glu His Pro Arg Phe Gly Ser Thr
      210            215            220
Ile Glu Gly Leu Ala Lys Leu Lys Pro Ala Phe Lys Lys Asp Gly Thr
      225            230            235            240
Val Thr Ala Gly Asn Ala Ser Gly Leu Asn Asp Cys Ala Ala Val Leu
      245            250            255
Val Ile Met Ser Ala Glu Lys Ala Lys Glu Leu Gly Val Lys Pro Leu
      260            265            270
Ala Lys Ile Val Ser Tyr Gly Ser Ala Gly Val Asp Pro Ala Ile Met
      275            280            285
Gly Tyr Gly Pro Phe Tyr Ala Thr Lys Ala Ala Ile Glu Lys Ala Gly
      290            295            300
Trp Thr Val Asp Glu Leu Asp Leu Ile Glu Ser Asn Glu Ala Phe Ala
      305            310            315            320
Ala Gln Ser Leu Ala Val Ala Lys Asp Leu Lys Phe Asp Met Asn Lys
      325            330            335

```

[illegible]

```
<210> SEQ ID NO 3
<211> LENGTH: 1179
<212> TYPE: DNA
<213> ORGANISM: Clostridium acetobutylicum
```

<400> SEQUENCE: 3

atgagagatg	tagtaatatg	aagtgcctga	agaactgcaa	taggagcata	tggaaaaaca	60
ttaaaggatg	tacctgcaac	agagttagga	gctatagtaa	taaaggaagc	tgtagaaga	120
gctaataata	atccaaatga	gattaatgaa	gttatttttg	gaaatgtact	tcaagctgga	180
ttaggccaaa	accagcaag	acaagcagca	gtaaaagcag	gattaccttt	agaaacacct	240
gcgtttacaa	tcaataaggt	ttgtggttca	ggtttaagat	ctataagttt	agcagctcaa	300
attataaaag	ctggagatgc	tgataccatt	gtagtagggt	gtatggaaaa	tatgtctaga	360
tcaccatatt	tgattaacaa	tcagagatgg	ggccaagaa	tgggagatag	tgaattagtt	420
gatgaaatga	taaaggatgg	tttgtgggat	gcatttaatg	gatatcatat	gggagtaact	480
gcagaaaaata	ttgcagaaca	atggaatata	acaagagaag	agcaagatga	attttcactt	540
atgtcacaac	aaaaagctga	aaaagccatt	aaaaatggag	aatttaagga	tgaatatagtt	600
cctgtattaa	taaagactaa	aaaaggtgaa	atagtctttg	atcaagatga	atttcctaga	660
ttcggaacaa	ctattgaagc	attaagaaaa	cttaaaccta	ttttcaagga	aaatggtact	720
gttacagcag	gtaatgcctc	cggattaaat	gatggagctg	cagcactagt	aataatgagc	780
gctgataaag	ctaacgctct	cggaaataaa	ccacttgcta	agattacttc	ttacggatca	840
tatggggtag	atccatcaat	aatgggatat	ggagcttttt	atgcaactaa	agctgcctta	900
gataaaatta	atttaaaacc	tgaagactta	gattttaattg	aagctaacga	ggcatatgct	960
tctcaaagta	tagcagtaac	tagagattta	aatttagata	tgagtaaagt	taatgttaat	1020
ggtggagcta	tagcacttgg	acatccaata	ggtgcactcg	gtgcacgtat	tttagtaaca	1080
ttactatacg	ctatgcaaaa	aagagattca	aaaaaaggtc	ttgctactct	atgtattggg	1140
ggaggtcagc	gaacagctct	cgtagtgtgaa	agaqactaa			1179

```
<210> SEQ ID NO 4
<211> LENGTH: 392
<212> TYPE: PRT
<213> ORGANISM: Clostridium acetobutylicum
```

<400> SEQUENCE: 4

Met	Arg	Asp	Val	Val	Ile	Val	Ser	Ala	Val	Arg	Thr	Ala	Ile	Gly	Ala
1				5					10					15	
Tyr	Gly	Lys	Thr	Leu	Lys	Asp	Val	Pro	Ala	Thr	Glu	Leu	Gly	Ala	Ile
			20					25					30		
Val	Ile	Lys	Glu	Ala	Val	Arg	Arg	Ala	Asn	Ile	Asn	Pro	Asn	Glu	Ile
		35					40					45			
Asn	Glu	Val	Ile	Phe	Gly	Asn	Val	Leu	Gln	Ala	Gly	Leu	Gly	Gln	Asn
	50					55					60				

-continued

Pro Ala Arg Gln Ala Ala Val Lys Ala Gly Leu Pro Leu Glu Thr Pro
 65 70 75 80
 Ala Phe Thr Ile Asn Lys Val Cys Gly Ser Gly Leu Arg Ser Ile Ser
 85 90 95
 Leu Ala Ala Gln Ile Ile Lys Ala Gly Asp Ala Asp Thr Ile Val Val
 100 105 110
 Gly Gly Met Glu Asn Met Ser Arg Ser Pro Tyr Leu Ile Asn Asn Gln
 115 120 125
 Arg Trp Gly Gln Arg Met Gly Asp Ser Glu Leu Val Asp Glu Met Ile
 130 135 140
 Lys Asp Gly Leu Trp Asp Ala Phe Asn Gly Tyr His Met Gly Val Thr
 145 150 155 160
 Ala Glu Asn Ile Ala Glu Gln Trp Asn Ile Thr Arg Glu Glu Gln Asp
 165 170 175
 Glu Phe Ser Leu Met Ser Gln Gln Lys Ala Glu Lys Ala Ile Lys Asn
 180 185 190
 Gly Glu Phe Lys Asp Glu Ile Val Pro Val Leu Ile Lys Thr Lys Lys
 195 200 205
 Gly Glu Ile Val Phe Asp Gln Asp Glu Phe Pro Arg Phe Gly Asn Thr
 210 215 220
 Ile Glu Ala Leu Arg Lys Leu Lys Pro Ile Phe Lys Glu Asn Gly Thr
 225 230 235 240
 Val Thr Ala Gly Asn Ala Ser Gly Leu Asn Asp Gly Ala Ala Ala Leu
 245 250 255
 Val Ile Met Ser Ala Asp Lys Ala Asn Ala Leu Gly Ile Lys Pro Leu
 260 265 270
 Ala Lys Ile Thr Ser Tyr Gly Ser Tyr Gly Val Asp Pro Ser Ile Met
 275 280 285
 Gly Tyr Gly Ala Phe Tyr Ala Thr Lys Ala Ala Leu Asp Lys Ile Asn
 290 295 300
 Leu Lys Pro Glu Asp Leu Asp Leu Ile Glu Ala Asn Glu Ala Tyr Ala
 305 310 315 320
 Ser Gln Ser Ile Ala Val Thr Arg Asp Leu Asn Leu Asp Met Ser Lys
 325 330 335
 Val Asn Val Asn Gly Gly Ala Ile Ala Leu Gly His Pro Ile Gly Ala
 340 345 350
 Ser Gly Ala Arg Ile Leu Val Thr Leu Leu Tyr Ala Met Gln Lys Arg
 355 360 365
 Asp Ser Lys Lys Gly Leu Ala Thr Leu Cys Ile Gly Gly Gly Gln Gly
 370 375 380
 Thr Ala Leu Val Val Glu Arg Asp
 385 390

<210> SEQ ID NO 5

<211> LENGTH: 849

<212> TYPE: DNA

<213> ORGANISM: Clostridium acetobutylicum

<400> SEQUENCE: 5

atgaaaaagg tatgtgttat aggtgcaggt actatggggt caggaattgc tcaggcattt	60
gcagctaaag gatttgaagt agtattaaga gatattaaag atgaatttgt tgatagagga	120
ttagatttta tcaataaaaa tctttctaaa ttagttaaaa aaggaaagat agaagaagct	180
actaaagttg aaatcttaac tagaatttcc ggaacagttg accttaatat ggcagctgat	240

-continued

```

tgcgatttag ttatagaagc agctgttgaa agaattggata ttaaaaagca gatttttgct 300
gacttagaca atatatgcaa gccagaaaca attcttgcac caaatacatc atcactttca 360
ataacagaag tggcatcagc aactaaaaga cctgataagg ttataggatg gcattttctt 420
aatccagctc ctgttatgaa gctttagag gtaataagag gaatagctac atcacaagaa 480
acttttgatg cagttaaaga gacatctata gcaataggaa aagatcctgt agaagtagca 540
gaagcaccag gatttgttgt aaatagaata ttaataccaa tgattaatga agcagttggt 600
atattagcag aaggaatagc ttcagtagaa gacatagata aagctatgaa acttgagact 660
aatcacccaa tgggaccatt agaattaggt gattttatag gtcttgatat atgtcttgct 720
ataatggatg tttatactc agaaactgga gattctaagt atagaccaca tacattactt 780
aagaagtatg taagagcagg atggcttga agaaaatcag gaaaagggtt ctacgattat 840
tcaaaataa 849

```

<210> SEQ ID NO 6

<211> LENGTH: 282

<212> TYPE: PRT

<213> ORGANISM: Clostridium acetobutylicum

<400> SEQUENCE: 6

```

Met Lys Lys Val Cys Val Ile Gly Ala Gly Thr Met Gly Ser Gly Ile
1           5           10          15
Ala Gln Ala Phe Ala Ala Lys Gly Phe Glu Val Val Leu Arg Asp Ile
20          25          30
Lys Asp Glu Phe Val Asp Arg Gly Leu Asp Phe Ile Asn Lys Asn Leu
35          40          45
Ser Lys Leu Val Lys Lys Gly Lys Ile Glu Glu Ala Thr Lys Val Glu
50          55          60
Ile Leu Thr Arg Ile Ser Gly Thr Val Asp Leu Asn Met Ala Ala Asp
65          70          75          80
Cys Asp Leu Val Ile Glu Ala Ala Val Glu Arg Met Asp Ile Lys Lys
85          90          95
Gln Ile Phe Ala Asp Leu Asp Asn Ile Cys Lys Pro Glu Thr Ile Leu
100         105         110
Ala Ser Asn Thr Ser Ser Leu Ser Ile Thr Glu Val Ala Ser Ala Thr
115         120         125
Lys Arg Pro Asp Lys Val Ile Gly Met His Phe Phe Asn Pro Ala Pro
130         135         140
Val Met Lys Leu Val Glu Val Ile Arg Gly Ile Ala Thr Ser Gln Glu
145         150         155         160
Thr Phe Asp Ala Val Lys Glu Thr Ser Ile Ala Ile Gly Lys Asp Pro
165         170         175
Val Glu Val Ala Glu Ala Pro Gly Phe Val Val Asn Arg Ile Leu Ile
180         185         190
Pro Met Ile Asn Glu Ala Val Gly Ile Leu Ala Glu Gly Ile Ala Ser
195         200         205
Val Glu Asp Ile Asp Lys Ala Met Lys Leu Gly Ala Asn His Pro Met
210         215         220
Gly Pro Leu Glu Leu Gly Asp Phe Ile Gly Leu Asp Ile Cys Leu Ala
225         230         235         240
Ile Met Asp Val Leu Tyr Ser Glu Thr Gly Asp Ser Lys Tyr Arg Pro
245         250         255

```

His Thr Leu Leu Lys Lys Tyr Val Arg Ala Gly Trp Leu Gly Arg Lys
260 265 270

Ser Gly Lys Gly Phe Tyr Asp Tyr Ser Lys
275 280

```
<210> SEQ ID NO 7
<211> LENGTH: 786
<212> TYPE: DNA
<213> ORGANISM: Clostridium acetobutylicum
```

<400> SEQUENCE: 7

atggaactaa	acaatgtcat	ccttgaaaag	gaaggtaaag	ttgctgtagt	taccattaac	60
agacctaag	cattaaatgc	gttaaatagt	gatacactaa	aagaaatgga	ttatgttata	120
ggtgaaattg	aaaatgatag	cgaagtactt	gcagtaattt	taactggagc	aggagaaaaa	180
tcatttgtag	caggagcaga	tatttctgag	atgaaggaaa	tgaataccat	tgaaggtaga	240
aaattcggga	tacttggaag	taaagtgttt	agaagattag	aacttcttga	aaagcctgta	300
atagcagctg	ttaatggttt	tgctttagga	ggcggatcgc	aaatagctat	gtcttgtgat	360
ataagaatag	cttcaagcaa	cgcaagattt	ggtaaccag	aagtaggtct	cggaataaca	420
cctgggtttg	gtggtacaca	aagactttca	agattagtgt	gaatgggcat	ggcaaagcag	480
cttatattta	ctgcacaaaa	tataaaggca	gatgaagcat	taagaatcgg	acttgtaaat	540
aaggtagtag	aacctagtga	attaatgaat	acagcaaaag	aaattgcaaa	caaaattgtg	600
agcaatgctc	cagtagctgt	taagttaagc	aaacaggcta	ttaatagagg	aatgcagtgt	660
gatattgata	ctgcttttagc	atttgaatca	gaagcatttg	gagaatgctt	ttcaacagag	720
gatcaaaagg	atgcaatgac	agctttcata	gagaaaaagaa	aaattgaagg	cttcaaaaat	780
agataq						786

```
<210> SEQ ID NO 8
<211> LENGTH: 261
<212> TYPE: PRT
<213> ORGANISM: Clostridium acetobutylicum
```

<400> SEQUENCE: 8

Met 1	Glu	Leu	Asn	Asn 5	Val	Ile	Leu	Glu	Lys 10	Glu	Gly	Lys	Val	Ala 15	Val
Val	Thr	Ile	Asn 20	Arg	Pro	Lys	Ala	Leu 25	Asn	Ala	Leu	Asn 30	Ser	Asp	Thr
Leu	Lys	Glu	Met	Asp	Tyr	Val	Ile 40	Gly	Glu	Ile	Glu	Asn 45	Asp	Ser	Glu
Val	Leu 50	Ala	Val	Ile	Leu	Thr 55	Gly	Ala	Gly	Glu	Lys 60	Ser	Phe	Val	Ala
Gly 65	Ala	Asp	Ile	Ser	Glu 70	Met	Lys	Glu	Met	Asn 75	Thr	Ile	Glu	Gly	Arg 80
Lys	Phe	Gly	Ile	Leu 85	Gly	Asn	Lys	Val	Phe	Arg 90	Arg	Leu	Glu	Leu 95	Leu
Glu	Lys	Pro	Val 100	Ile	Ala	Ala	Val	Asn 105	Gly	Phe	Ala	Leu	Gly 110	Gly	Gly
Cys	Glu	Ile	Ala 115	Met	Ser	Cys	Asp 120	Ile	Arg	Ile	Ala	Ser 125	Ser	Asn	Ala
Arg	Phe 130	Gly	Gln	Pro	Glu	Val 135	Gly	Leu	Gly	Ile	Thr 140	Pro	Gly	Phe	Gly
Gly 145	Thr	Gln	Arg	Leu 150	Ser	Arg	Leu	Val	Gly	Met 155	Gly	Met	Ala	Lys	Gln 160

-continued

Leu Ile Phe Thr Ala Gln Asn Ile Lys Ala Asp Glu Ala Leu Arg Ile
 165 170 175
 Gly Leu Val Asn Lys Val Val Glu Pro Ser Glu Leu Met Asn Thr Ala
 180 185 190
 Lys Glu Ile Ala Asn Lys Ile Val Ser Asn Ala Pro Val Ala Val Lys
 195 200 205
 Leu Ser Lys Gln Ala Ile Asn Arg Gly Met Gln Cys Asp Ile Asp Thr
 210 215 220
 Ala Leu Ala Phe Glu Ser Glu Ala Phe Gly Glu Cys Phe Ser Thr Glu
 225 230 235 240
 Asp Gln Lys Asp Ala Met Thr Ala Phe Ile Glu Lys Arg Lys Ile Glu
 245 250 255
 Gly Phe Lys Asn Arg
 260

<210> SEQ ID NO 9
 <211> LENGTH: 1197
 <212> TYPE: DNA
 <213> ORGANISM: Clostridium acetobutylicum

<400> SEQUENCE: 9

```

atgatagtaa aagcaaagtt tgtaaaagga tttatcagag atgtacatcc ttatggttgc      60
agaaggggaag tactaaatca aatagattat tgtaagaagg ctattggggtt tagggggacca    120
aagaaggttt taattgttgg agcctcatct gggtttggtc ttgctactag aatttcagtt    180
gcatttggag gtccagaagc tcacacaatt ggagtatcct atgaaacagg agctacagat    240
agaagaatag gaacagcggg atggtataat aacatatttt ttaaagaatt tgctaaaaaa    300
aaaggattag ttgcaaaaaa cttcattgag gatgcctttt ctaatgaaac caagataaaa    360
gttattaagt atataaagga tgaatttggg aaaatagatt tatttgttta tagtttagct    420
gcgcctagga gaaaggacta taaaactgga aatgtttata cttcaagaat aaaaacaatt    480
ttaggagatt ttgagggacc gactattgat gttgaaagag acgagattac tttaaaaaag    540
gttagtagtg ctagcattga agaaattgaa gaaactagaa aggtaatggg tggagaggat    600
tggcaagagt ggtgtgaaga gctgctttat gaagattggt tttcggataa agcaactacc    660
atagcatact cgtatatagg atccccaaga acctacaaga tatatagaga aggtactata    720
ggaatagcta aaaaggatct tgaagataag gctaagctta taaatgaaaa acttaacaga    780
gttataggtg gtagagcctt tgtgtctgtg aataaagcat tagttacaaa agcaagtgca    840
tatattccaa cttttcctct ttatgcagct attttatata aggtcatgaa agaaaaaaat    900
attcatgaaa attgtattat gcaaattgag agaattgttt ctgaaaaaat atattcaaat    960
gaaaaaatac aatttgatga caagggaaga ttaaggatgg acgatttaga gcttagaaaa   1020
gacgttcaag acgaagtga tagaatatgg agtaatatta ctctgaaaa ttttaaggaa   1080
ttatctgatt ataagggata caaaaaagaa ttcattgaact taaacggttt tgatctagat   1140
ggggttgatt atagtaaaga cctggatata gaattattaa gaaaattaga accttaa     1197
  
```

<210> SEQ ID NO 10
 <211> LENGTH: 398
 <212> TYPE: PRT
 <213> ORGANISM: Clostridium acetobutylicum

<400> SEQUENCE: 10

Met 1	Ile	Val	Lys 5	Ala	Lys	Phe	Val	Lys 10	Gly	Phe	Ile	Arg	Asp	Val 15	His
Pro	Tyr	Gly	Cys 20	Arg	Arg	Glu	Val	Leu 25	Asn	Gln	Ile	Asp	Tyr 30	Cys	Lys
Lys	Ala	Ile	Gly 35	Phe	Arg	Gly	Pro 40	Lys	Lys	Val	Leu 45	Ile	Val	Gly	Ala
Ser 50	Ser	Gly	Phe	Gly	Leu	Ala 55	Thr	Arg	Ile	Ser	Val 60	Ala	Phe	Gly	Gly
Pro 65	Glu	Ala	His	Thr	Ile 70	Gly	Val	Ser	Tyr	Glu 75	Thr	Gly	Ala	Thr	Asp 80
Arg	Arg	Ile	Gly	Thr 85	Ala	Gly	Trp	Tyr	Asn 90	Asn	Ile	Phe	Phe	Lys 95	Glu
Phe	Ala	Lys	Lys 100	Lys	Gly	Leu	Val	Ala 105	Lys	Asn	Phe	Ile	Glu 110	Asp	Ala
Phe	Ser	Asn	Glu 115	Thr	Lys	Asp	Lys 120	Val	Ile	Lys	Tyr	Ile 125	Lys	Asp	Glu
Phe 130	Gly	Lys	Ile	Asp	Leu	Phe 135	Val	Tyr	Ser	Leu 140	Ala	Ala	Pro	Arg	Arg
Lys 145	Asp	Tyr	Lys	Thr	Gly 150	Asn	Val	Tyr	Thr	Ser 155	Arg	Ile	Lys	Thr	Ile 160
Leu	Gly	Asp	Phe	Glu 165	Gly	Pro	Thr	Ile	Asp 170	Val	Glu	Arg	Asp	Glu 175	Ile
Thr	Leu	Lys	Lys 180	Val	Ser	Ser	Ala	Ser 185	Ile	Glu	Glu	Ile	Glu 190	Glu	Thr
Arg	Lys	Val 195	Met	Gly	Gly	Glu	Asp 200	Trp	Gln	Glu	Trp	Cys 205	Glu	Glu	Leu
Leu 210	Tyr	Glu	Asp	Cys	Phe	Ser 215	Asp	Lys	Ala	Thr 220	Thr	Ile	Ala	Tyr	Ser
Tyr 225	Ile	Gly	Ser	Pro	Arg 230	Thr	Tyr	Lys	Ile	Tyr 235	Arg	Glu	Gly	Thr	Ile 240
Gly	Ile	Ala	Lys 245	Lys	Asp	Leu	Glu	Asp	Lys 250	Ala	Lys	Leu	Ile	Asn 255	Glu
Lys	Leu	Asn	Arg 260	Val	Ile	Gly	Gly	Arg 265	Ala	Phe	Val	Ser	Val 270	Asn	Lys
Ala	Leu	Val 275	Thr	Lys	Ala	Ser	Ala 280	Tyr	Ile	Pro	Thr	Phe 285	Pro	Leu	Tyr
Ala 290	Ala	Ile	Leu	Tyr	Lys 295	Val	Met	Lys	Glu	Lys 300	Asn	Ile	His	Glu	Asn
Cys 305	Ile	Met	Gln	Ile	Glu 310	Arg	Met	Phe	Ser	Glu 315	Lys	Ile	Tyr	Ser	Asn 320
Glu	Lys	Ile	Gln 325	Phe	Asp	Asp	Lys	Gly	Arg 330	Leu	Arg	Met	Asp	Asp 335	Leu
Glu	Leu	Arg	Lys 340	Asp	Val	Gln	Asp	Glu 345	Val	Asp	Arg	Ile	Trp 350	Ser	Asn
Ile	Thr 355	Pro	Glu	Asn	Phe	Lys 360	Glu	Leu	Ser	Asp	Tyr	Lys 365	Gly	Tyr	Lys
Lys 370	Glu	Phe	Met	Asn	Leu 375	Asn	Gly	Phe	Asp	Leu 380	Asp	Gly	Val	Asp	Tyr
Ser 385	Lys	Asp	Leu	Asp 390	Ile	Glu	Leu	Leu	Arg 395	Lys	Leu	Glu	Pro		

<210> SEQ ID NO 11
<211> LENGTH: 1407
<212> TYPE: DNA

-continued

<213> ORGANISM: Clostridium beijerinckii

<400> SEQUENCE: 11

```

atgaataaag acacactaat acctacaact aaagatttaa aagtaaaaac aaatggtgaa    60
aacattaatt taaagaacta caaggataat tcttcattgt tcggagtatt cgaaaatgtt    120
gaaaatgcta taagcagcgc tgtacacgca caaaagatat tatcccttca ttatacaaaa    180
gagcaaagag aaaaaatcat aactgagata agaaaggccg cattacaaaa taaagaggtc    240
ttggctacaa tgattctaga agaaacacat atgggaagat atgaggataa aatattaaaa    300
catgaattgg tagctaaata tactcctggt acagaagatt taactactac tgcttggtca    360
ggtgataatg gtcttacagt tgtagaaatg tctccatatg gtgttatagg tgcaataact    420
ccttctacga atccaactga aactgtaata tgtaatagca taggcattgat agctgctgga    480
aatgctgtag tatttaacgg accccattgc gctaaaaaat gtgttgcttt tgctgttgaa    540
atgataaata aggcaattat ttcattgtgc ggtcctgaaa atctagtaac aactataaaa    600
aatccaacta tggagtctct agatgcaatt attaagcatc cttcaataaa acttctttgc    660
ggaactgggg gtccaggaat ggtaaaaacc ctcttaaatt ctggtaagaa agctataggt    720
gctggtgctg gaaatccacc agttattgta gatgatactg ctgatataga aaaggctggt    780
aggagcatca ttgaaggctg ttcttttgat aataatttac cttgtattgc agaaaaagaa    840
gtatttgttt ttgagaatgt tgcagatgat ttaatatcta acatgctaaa aaataatgct    900
gtaattataa atgaagatca agtatcaaaa ttaatagatt tagtattaca aaaaaataat    960
gaaactcaag aatactttat aaacaaaaaa tgggtaggaa aagatgcaaa attattctta   1020
gatgaaatag atgttgagtc tccttcaa atgttaaatgca taatctgcga agtaaatgca   1080
aatcatccat ttgttatgac agaactcatg atgccaatat tgccaattgt aagagttaaa   1140
gatatagatg aagctattaa atatgcaaag atagcagaac aaaatagaaa acatagtgcc   1200
tatattttatt ctaaaaatat agacaaccta aatagatttg aaagagaaat agatactact   1260
atttttgtaa agaatgctaa atcttttgct ggtgttggtt atgaagcaga aggatttaca   1320
actttcacta ttgctggatc tactggtgag ggaataacct ctgcaaggaa tttacaaga   1380
caaagaagat gtgtacttgc cggtctaa                                     1407

```

<210> SEQ ID NO 12

<211> LENGTH: 468

<212> TYPE: PRT

<213> ORGANISM: Clostridium beijerinckii

<400> SEQUENCE: 12

```

Met Asn Lys Asp Thr Leu Ile Pro Thr Thr Lys Asp Leu Lys Val Lys
1          5          10          15

Thr Asn Gly Glu Asn Ile Asn Leu Lys Asn Tyr Lys Asp Asn Ser Ser
20          25          30

Cys Phe Gly Val Phe Glu Asn Val Glu Asn Ala Ile Ser Ser Ala Val
35          40          45

His Ala Gln Lys Ile Leu Ser Leu His Tyr Thr Lys Glu Gln Arg Glu
50          55          60

Lys Ile Ile Thr Glu Ile Arg Lys Ala Ala Leu Gln Asn Lys Glu Val
65          70          75          80

Leu Ala Thr Met Ile Leu Glu Glu Thr His Met Gly Arg Tyr Glu Asp
85          90          95

```

-continued

Lys Ile Leu Lys His Glu Leu Val Ala Lys Tyr Thr Pro Gly Thr Glu
 100 105 110
 Asp Leu Thr Thr Thr Ala Trp Ser Gly Asp Asn Gly Leu Thr Val Val
 115 120 125
 Glu Met Ser Pro Tyr Gly Val Ile Gly Ala Ile Thr Pro Ser Thr Asn
 130 135 140
 Pro Thr Glu Thr Val Ile Cys Asn Ser Ile Gly Met Ile Ala Ala Gly
 145 150 155 160
 Asn Ala Val Val Phe Asn Gly His Pro Cys Ala Lys Lys Cys Val Ala
 165 170 175
 Phe Ala Val Glu Met Ile Asn Lys Ala Ile Ile Ser Cys Gly Gly Pro
 180 185 190
 Glu Asn Leu Val Thr Thr Ile Lys Asn Pro Thr Met Glu Ser Leu Asp
 195 200 205
 Ala Ile Ile Lys His Pro Ser Ile Lys Leu Leu Cys Gly Thr Gly Gly
 210 215 220
 Pro Gly Met Val Lys Thr Leu Leu Asn Ser Gly Lys Lys Ala Ile Gly
 225 230 235 240
 Ala Gly Ala Gly Asn Pro Pro Val Ile Val Asp Asp Thr Ala Asp Ile
 245 250 255
 Glu Lys Ala Gly Arg Ser Ile Ile Glu Gly Cys Ser Phe Asp Asn Asn
 260 265 270
 Leu Pro Cys Ile Ala Glu Lys Glu Val Phe Val Phe Glu Asn Val Ala
 275 280 285
 Asp Asp Leu Ile Ser Asn Met Leu Lys Asn Asn Ala Val Ile Ile Asn
 290 295 300
 Glu Asp Gln Val Ser Lys Leu Ile Asp Leu Val Leu Gln Lys Asn Asn
 305 310 315 320
 Glu Thr Gln Glu Tyr Phe Ile Asn Lys Lys Trp Val Gly Lys Asp Ala
 325 330 335
 Lys Leu Phe Leu Asp Glu Ile Asp Val Glu Ser Pro Ser Asn Val Lys
 340 345 350
 Cys Ile Ile Cys Glu Val Asn Ala Asn His Pro Phe Val Met Thr Glu
 355 360 365
 Leu Met Met Pro Ile Leu Pro Ile Val Arg Val Lys Asp Ile Asp Glu
 370 375 380
 Ala Ile Lys Tyr Ala Lys Ile Ala Glu Gln Asn Arg Lys His Ser Ala
 385 390 395 400
 Tyr Ile Tyr Ser Lys Asn Ile Asp Asn Leu Asn Arg Phe Glu Arg Glu
 405 410 415
 Ile Asp Thr Thr Ile Phe Val Lys Asn Ala Lys Ser Phe Ala Gly Val
 420 425 430
 Gly Tyr Glu Ala Glu Gly Phe Thr Thr Phe Thr Ile Ala Gly Ser Thr
 435 440 445
 Gly Glu Gly Ile Thr Ser Ala Arg Asn Phe Thr Arg Gln Arg Arg Cys
 450 455 460
 Val Leu Ala Gly
 465

<210> SEQ ID NO 13

<211> LENGTH: 1215

<212> TYPE: DNA

<213> ORGANISM: Clostridium acetobutylicum

<400> SEQUENCE: 13

-continued

```

atggttgatt tcgaatatc aataccaact agaatttttt tcggtaaaga taagataaat    60
gtacttgga gagagcttaa aaaatatggt tctaaagtgc ttatagttaa tggaggagga    120
agtataaaga gaaatggaat atatgataaa gctgtaagta tacttgaaaa aaacagtatt    180
aaattttatg aacttgagg agtagagcca aatccaagag taactacagt tgaaaaagga    240
gttaaaatat gtagagaaaa tggagtgaa gtagtactag ctataggagg aggaagtgc    300
atagattgag caaagggtat agcagcagca tgtgaatatg atggaaatcc atgggatatt    360
gtgttagatg gctcaaaaat aaaaagggtg ctctctatag ctagtatatt aaccattgct    420
gcaacaggat cagaaatgga tacgtgggca gtaataaata atatggatac aaacgaaaaa    480
ctaattgcgg cacatccaga tatggctcct aagtttttcta tattagatcc aacgtatacg    540
tataccgtac ctaccaatca aacagcagca ggaacagctg atattatgag tcatatattt    600
gagggtgatt ttagtaatac aaaaacagca tatttgcagg atagaatggc agaagcgtaa    660
ttaagaactt gtattaaata tggaggaata gctcttgaga agccggatga ttatgaggca    720
agagccaatc taatgtgggc ttcaagtctt gcgataaatg gacttttaac atatggtaaa    780
gacactaatt ggagtgtaca cttaatggaa catgaattaa gtgcttatta cgacataaca    840
cacggcgtag ggcttgcaat ttaaacacct aattggatgg agtatatttt aaataatgat    900
acagtgtaca agtttgttga atatggtgta aatgtttggg gaatagacaa agaaaaaat    960
cactatgaca tagcacatca agcaatacaa aaaacaagag attactttgt aaatgtacta   1020
ggttttacat ctagactgag agatgttgga attgaagaag aaaaattgga cataatggca   1080
aaggaatcag taaagcttac aggaggaacc ataggaaacc taagaccagt aaacgcctcc   1140
gaagtcctac aaatattcaa aaaatctgtg taaaacgcct ccgaagtcct acaaattatc   1200
aaaaaatctg tgtaa                                           1215

```

```

<210> SEQ ID NO 14
<211> LENGTH: 390
<212> TYPE: PRT
<213> ORGANISM: Clostridium acetobutylicum

```

```

<400> SEQUENCE: 14

```

```

Met Val Asp Phe Glu Tyr Ser Ile Pro Thr Arg Ile Phe Phe Gly Lys
1           5           10          15

Asp Lys Ile Asn Val Leu Gly Arg Glu Leu Lys Lys Tyr Gly Ser Lys
20          25          30

Val Leu Ile Val Tyr Gly Gly Gly Ser Ile Lys Arg Asn Gly Ile Tyr
35          40          45

Asp Lys Ala Val Ser Ile Leu Glu Lys Asn Ser Ile Lys Phe Tyr Glu
50          55          60

Leu Ala Gly Val Glu Pro Asn Pro Arg Val Thr Thr Val Glu Lys Gly
65          70          75          80

Val Lys Ile Cys Arg Glu Asn Gly Val Glu Val Val Leu Ala Ile Gly
85          90          95

Gly Gly Ser Ala Ile Asp Cys Ala Lys Val Ile Ala Ala Ala Cys Glu
100         105         110

Tyr Asp Gly Asn Pro Trp Asp Ile Val Leu Asp Gly Ser Lys Ile Lys
115         120         125

Arg Val Leu Pro Ile Ala Ser Ile Leu Thr Ile Ala Ala Thr Gly Ser
130         135         140

```

-continued

Glu Met Asp Thr Trp Ala Val Ile Asn Asn Met Asp Thr Asn Glu Lys
 145 150 155 160
 Leu Ile Ala Ala His Pro Asp Met Ala Pro Lys Phe Ser Ile Leu Asp
 165 170 175
 Pro Thr Tyr Thr Tyr Thr Val Pro Thr Asn Gln Thr Ala Ala Gly Thr
 180 185 190
 Ala Asp Ile Met Ser His Ile Phe Glu Val Tyr Phe Ser Asn Thr Lys
 195 200 205
 Thr Ala Tyr Leu Gln Asp Arg Met Ala Glu Ala Leu Leu Arg Thr Cys
 210 215 220
 Ile Lys Tyr Gly Gly Ile Ala Leu Glu Lys Pro Asp Asp Tyr Glu Ala
 225 230 235 240
 Arg Ala Asn Leu Met Trp Ala Ser Ser Leu Ala Ile Asn Gly Leu Leu
 245 250 255
 Thr Tyr Gly Lys Asp Thr Asn Trp Ser Val His Leu Met Glu His Glu
 260 265 270
 Leu Ser Ala Tyr Tyr Asp Ile Thr His Gly Val Gly Leu Ala Ile Leu
 275 280 285
 Thr Pro Asn Trp Met Glu Tyr Ile Leu Asn Asn Asp Thr Val Tyr Lys
 290 295 300
 Phe Val Glu Tyr Gly Val Asn Val Trp Gly Ile Asp Lys Glu Lys Asn
 305 310 315 320
 His Tyr Asp Ile Ala His Gln Ala Ile Gln Lys Thr Arg Asp Tyr Phe
 325 330 335
 Val Asn Val Leu Gly Leu Pro Ser Arg Leu Arg Asp Val Gly Ile Glu
 340 345 350
 Glu Glu Lys Leu Asp Ile Met Ala Lys Glu Ser Val Lys Leu Thr Gly
 355 360 365
 Gly Thr Ile Gly Asn Leu Arg Pro Val Asn Ala Ser Glu Val Leu Gln
 370 375 380
 Ile Phe Lys Lys Ser Val
 385 390

<210> SEQ ID NO 15
 <211> LENGTH: 1170
 <212> TYPE: DNA
 <213> ORGANISM: Clostridium acetobutylicum

<400> SEQUENCE: 15

atgctaagtt ttgattattc aataccaact aaagtttttt ttggaaaagg aaaaatagac	60
gtaattggag aagaaattaa gaaatatggc tcaagagtgct ttatagttta tggcggagga	120
agtataaaaa ggaacggtat atatgataga gcaacagcta tattaaaaga aaacaatata	180
gctttctatg aactttcagg agtagagcca aatcctagga taacaacagt aaaaaaggc	240
atagaaatat gtatagaaaa taatgtggat ttagtattag caataggggg aggaagtgca	300
atagactgtt ctaaggaat tgcagctgga gtttattatg atggcgatac atgggacatg	360
gttaaagatc catctaaat aactaaagtt cttccaattg caagtatact tactctttca	420
gcaacagggt ctgaaatgga tcaaatgca gtaatttcaa atatggagac taatgaaaag	480
cttgagtag gacatgatga tatgagacct aaattttcag tgtagatcc tacatatact	540
tttacagtac ctaaaaaatca aacagcagcg ggaacagctg acattatgag tcacaccttt	600
gaatcttact ttagtggtgt tgaagtgct tatgtgcagg acggtatagc agaagcaatc	660
ttaagaacat gtataaagta tggaaaaata gcaatggaga agactgatga ttacgaggct	720

-continued

```

agagctaatt tgatgtgggc ttcaagttaa gctataaatg gtctattatc acttggttaag 780
gatagaaaat ggagttgtca tcctatggaa cacgagttaa gtgcatatta tgatataaca 840
catggtgtag gacttgcaat tttaacacct aattggatgg aatatattct aaatgacgat 900
acacttcata aatttgtttc ttatggaata aatgtttggg gaatagacaa gaacaaagat 960
aactatgaaa tagcacgaga ggctattaaa aatacgagag aatactttaa ttcattgggt 1020
attccttcaa agcttagaga agttggaata ggaaaagata aactagaact aatggcaaag 1080
caagctgtta gaaattctgg aggaacaata ggaagttaa gaccaataaa tgcagaggat 1140
gttcttgaga tatttaaaaa atcttattaa 1170

```

<210> SEQ ID NO 16

<211> LENGTH: 389

<212> TYPE: PRT

<213> ORGANISM: Clostridium acetobutylicum

<400> SEQUENCE: 16

```

Met Leu Ser Phe Asp Tyr Ser Ile Pro Thr Lys Val Phe Phe Gly Lys
1           5           10          15
Gly Lys Ile Asp Val Ile Gly Glu Glu Ile Lys Lys Tyr Gly Ser Arg
20          25          30
Val Leu Ile Val Tyr Gly Gly Gly Ser Ile Lys Arg Asn Gly Ile Tyr
35          40          45
Asp Arg Ala Thr Ala Ile Leu Lys Glu Asn Asn Ile Ala Phe Tyr Glu
50          55          60
Leu Ser Gly Val Glu Pro Asn Pro Arg Ile Thr Thr Val Lys Lys Gly
65          70          75          80
Ile Glu Ile Cys Arg Glu Asn Asn Val Asp Leu Val Leu Ala Ile Gly
85          90          95
Gly Gly Ser Ala Ile Asp Cys Ser Lys Val Ile Ala Ala Gly Val Tyr
100         105         110
Tyr Asp Gly Asp Thr Trp Asp Met Val Lys Asp Pro Ser Lys Ile Thr
115        120        125
Lys Val Leu Pro Ile Ala Ser Ile Leu Thr Leu Ser Ala Thr Gly Ser
130        135        140
Glu Met Asp Gln Ile Ala Val Ile Ser Asn Met Glu Thr Asn Glu Lys
145        150        155        160
Leu Gly Val Gly His Asp Asp Met Arg Pro Lys Phe Ser Val Leu Asp
165        170        175
Pro Thr Tyr Thr Phe Thr Val Pro Lys Asn Gln Thr Ala Ala Gly Thr
180        185        190
Ala Asp Ile Met Ser His Thr Phe Glu Ser Tyr Phe Ser Gly Val Glu
195        200        205
Gly Ala Tyr Val Gln Asp Gly Ile Ala Glu Ala Ile Leu Arg Thr Cys
210        215        220
Ile Lys Tyr Gly Lys Ile Ala Met Glu Lys Thr Asp Asp Tyr Glu Ala
225        230        235        240
Arg Ala Asn Leu Met Trp Ala Ser Ser Leu Ala Ile Asn Gly Leu Leu
245        250        255
Ser Leu Gly Lys Asp Arg Lys Trp Ser Cys His Pro Met Glu His Glu
260        265        270
Leu Ser Ala Tyr Tyr Asp Ile Thr His Gly Val Gly Leu Ala Ile Leu
275        280        285

```

-continued

Thr Pro Asn Trp Met Glu Tyr Ile Leu Asn Asp Asp Thr Leu His Lys
 290 295 300
 Phe Val Ser Tyr Gly Ile Asn Val Trp Gly Ile Asp Lys Asn Lys Asp
 305 310 315 320
 Asn Tyr Glu Ile Ala Arg Glu Ala Ile Lys Asn Thr Arg Glu Tyr Phe
 325 330 335
 Asn Ser Leu Gly Ile Pro Ser Lys Leu Arg Glu Val Gly Ile Gly Lys
 340 345 350
 Asp Lys Leu Glu Leu Met Ala Lys Gln Ala Val Arg Asn Ser Gly Gly
 355 360 365
 Thr Ile Gly Ser Leu Arg Pro Ile Asn Ala Glu Asp Val Leu Glu Ile
 370 375 380
 Phe Lys Lys Ser Tyr
 385

<210> SEQ ID NO 17
 <211> LENGTH: 29
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Primer

<400> SEQUENCE: 17

caccatggaa ctaaacaatg tcaccccttg

29

<210> SEQ ID NO 18
 <211> LENGTH: 25
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Primer

<400> SEQUENCE: 18

cctccctatct atttttgaag ccttc

25

<210> SEQ ID NO 19
 <211> LENGTH: 28
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Primer

<400> SEQUENCE: 19

caccatgaaa aaggtatgtg ttataggt

28

<210> SEQ ID NO 20
 <211> LENGTH: 23
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Primer

<400> SEQUENCE: 20

catttgataa tggggattct tgt

23

<210> SEQ ID NO 21
 <211> LENGTH: 30
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Primer

<400> SEQUENCE: 21

-continued

caccatgaaa gaagttgtaa tagctagtgc 30

<210> SEQ ID NO 22
<211> LENGTH: 26
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer

<400> SEQUENCE: 22

ctagcacttt tctagcaata ttgctg 26

<210> SEQ ID NO 23
<211> LENGTH: 29
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer

<400> SEQUENCE: 23

caccatgcta agttttgatt attcaatac 29

<210> SEQ ID NO 24
<211> LENGTH: 25
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer

<400> SEQUENCE: 24

ttaataagat tttttaata tctca 25

<210> SEQ ID NO 25
<211> LENGTH: 30
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer

<400> SEQUENCE: 25

caccatgggt gatttcgaat attcaatacc 30

<210> SEQ ID NO 26
<211> LENGTH: 25
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer

<400> SEQUENCE: 26

ttacacagat tttttgaata ttgt 25

<210> SEQ ID NO 27
<211> LENGTH: 32
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer

<400> SEQUENCE: 27

caccatgaga gatgtagtaa tagtaagtgc tg 32

<210> SEQ ID NO 28
<211> LENGTH: 24
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence

-continued

<220> FEATURE:
<223> OTHER INFORMATION: Primer

<400> SEQUENCE: 28

ccgcaattgt atccatattg aacc 24

<210> SEQ ID NO 29
<211> LENGTH: 26
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer

<400> SEQUENCE: 29

caccatgata gtaaaagcaa agtttg 26

<210> SEQ ID NO 30
<211> LENGTH: 27
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer

<400> SEQUENCE: 30

gcttaaagct taaaaccgct tctggcg 27

<210> SEQ ID NO 31
<211> LENGTH: 27
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer

<400> SEQUENCE: 31

caccatgaat aaagacacac taatacc 27

<210> SEQ ID NO 32
<211> LENGTH: 24
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer

<400> SEQUENCE: 32

gccagaccat ctttgaaaat gcgc 24

<210> SEQ ID NO 33
<211> LENGTH: 35
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer

<400> SEQUENCE: 33

catgcatgca aaggagggtta gtagaatgaa agaag 35

<210> SEQ ID NO 34
<211> LENGTH: 38
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer

<400> SEQUENCE: 34

gtcctgcagg gcgcgcccaa tactttotag cacttttc 38

-continued

<210> SEQ ID NO 35
<211> LENGTH: 39
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer

<400> SEQUENCE: 35

catgtcgaca aaggaggtct gtttaatgaa aaaggtatg 39

<210> SEQ ID NO 36
<211> LENGTH: 28
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer

<400> SEQUENCE: 36

gtcgcatgcc ttgtaaactt attttgaa 28

<210> SEQ ID NO 37
<211> LENGTH: 44
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer

<400> SEQUENCE: 37

catagatctg gatccaaagg agggtaggga aatgatagta aaag 44

<210> SEQ ID NO 38
<211> LENGTH: 30
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer

<400> SEQUENCE: 38

catgtcgacg tgcagccttt ttaaggttct 30

<210> SEQ ID NO 39
<211> LENGTH: 38
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer

<400> SEQUENCE: 39

catgaattca cgcgtaaagg aggtattagt catggaac 38

<210> SEQ ID NO 40
<211> LENGTH: 30
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer

<400> SEQUENCE: 40

gtcggatccc ttacctccta tctatttttg 30

<210> SEQ ID NO 41
<211> LENGTH: 42
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer

-continued

<400> SEQUENCE: 41

catgcccggg ggtcaccaaa ggaggaatag ttcatgaata aa 42

<210> SEQ ID NO 42

<211> LENGTH: 32

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Primer

<400> SEQUENCE: 42

catggttaac aagaagttag ccggcaagta ca 32

<210> SEQ ID NO 43

<211> LENGTH: 41

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Primer

<400> SEQUENCE: 43

catggttaac aaaggagggg ttaaaatggt tgatttcgaa t 41

<210> SEQ ID NO 44

<211> LENGTH: 37

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Primer

<400> SEQUENCE: 44

catggcatgc gtttaaactg aggtttacac agatttt 37

<210> SEQ ID NO 45

<211> LENGTH: 17

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Primer

<400> SEQUENCE: 45

gtaaacgac ggccagt 17

<210> SEQ ID NO 46

<211> LENGTH: 16

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Primer

<400> SEQUENCE: 46

aacagctatg accatg 16

<210> SEQ ID NO 47

<211> LENGTH: 22

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Primer

<400> SEQUENCE: 47

gcaggagatg ctgacgtaat aa 22

<210> SEQ ID NO 48

-continued

<211> LENGTH: 24
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer

<400> SEQUENCE: 48

ccaacctgct ttttcaatag ctgc 24

<210> SEQ ID NO 49
<211> LENGTH: 21
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer

<400> SEQUENCE: 49

cagagatggg gtcaaagaat g 21

<210> SEQ ID NO 50
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer

<400> SEQUENCE: 50

gtggttttat tccgagagcg 20

<210> SEQ ID NO 51
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer

<400> SEQUENCE: 51

ggtctatact tagaatctcc 20

<210> SEQ ID NO 52
<211> LENGTH: 24
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer

<400> SEQUENCE: 52

cggaaacagtt gaccttaata tggc 24

<210> SEQ ID NO 53
<211> LENGTH: 22
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer

<400> SEQUENCE: 53

gcctcatctg ggtttggtct tg 22

<210> SEQ ID NO 54
<211> LENGTH: 28
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer

<400> SEQUENCE: 54

-continued

cgcctaggag aaaggactat aaaactgg 28

<210> SEQ ID NO 55
<211> LENGTH: 22
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer

<400> SEQUENCE: 55

cagagttata ggtggtagag cc 22

<210> SEQ ID NO 56
<211> LENGTH: 24
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer

<400> SEQUENCE: 56

ccatcccgct gttctctattc ttct 24

<210> SEQ ID NO 57
<211> LENGTH: 22
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer

<400> SEQUENCE: 57

ccaatcctct ccaccatta cc 22

<210> SEQ ID NO 58
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer

<400> SEQUENCE: 58

cgtccatcct taatcttccc 20

<210> SEQ ID NO 59
<211> LENGTH: 23
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer

<400> SEQUENCE: 59

ccaactatgg aatccctaga tgc 23

<210> SEQ ID NO 60
<211> LENGTH: 22
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer

<400> SEQUENCE: 60

gcatagtctg cgaagtaa at gc 22

<210> SEQ ID NO 61
<211> LENGTH: 24
<212> TYPE: DNA

-continued

<213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Primer

<400> SEQUENCE: 61

ggatctactg gtgaaggcat aacc 24

<210> SEQ ID NO 62
 <211> LENGTH: 23
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Primer

<400> SEQUENCE: 62

ggcatcatga gttctgtcat gac 23

<210> SEQ ID NO 63
 <211> LENGTH: 25
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Primer

<400> SEQUENCE: 63

gccttcaatg atactcttac cagcc 25

<210> SEQ ID NO 64
 <211> LENGTH: 22
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Primer

<400> SEQUENCE: 64

gcatttccag cagctatcat gc 22

<210> SEQ ID NO 65
 <211> LENGTH: 22
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Primer

<400> SEQUENCE: 65

ccttccata tgtgtttcct cc 22

<210> SEQ ID NO 66
 <211> LENGTH: 22
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Primer

<400> SEQUENCE: 66

gttgaagtag tactagctat ag 22

<210> SEQ ID NO 67
 <211> LENGTH: 22
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Primer

<400> SEQUENCE: 67

gacataacac acggcgtagg gc 22

-continued

<210> SEQ ID NO 68
<211> LENGTH: 22
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer

<400> SEQUENCE: 68

taagtgtaca ctccaattag tg 22

<210> SEQ ID NO 69
<211> LENGTH: 24
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer

<400> SEQUENCE: 69

gccatctaac acaatatccc atgg 24

<210> SEQ ID NO 70
<211> LENGTH: 24
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer

<400> SEQUENCE: 70

gcgatacatg ggacatgggt aaag 24

<210> SEQ ID NO 71
<211> LENGTH: 22
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer

<400> SEQUENCE: 71

tgcaactaac tcgtgttcca ta 22

<210> SEQ ID NO 72
<211> LENGTH: 21
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer

<400> SEQUENCE: 72

gttagccggc aagtacacat c 21

<210> SEQ ID NO 73
<211> LENGTH: 21
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer

<400> SEQUENCE: 73

actttctttc gctgtttca c 21

<210> SEQ ID NO 74
<211> LENGTH: 47
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:

-continued

<223> OTHER INFORMATION: Primer

<400> SEQUENCE: 74

catgaagctt ggcgcgccgg gacgcgtttt tgaaaaaat gaaaact 47

<210> SEQ ID NO 75

<211> LENGTH: 79

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Primer

<400> SEQUENCE: 75

catgaagctt gtttaaacctt ggtgaccttg aaaataatga aaacttatat tgttttgaaa 60

ataatgaaaa cttatatgtg 79

<210> SEQ ID NO 76

<211> LENGTH: 1197

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Codon optimized CAC0462 gene from Clostridium acetobutylicum

<400> SEQUENCE: 76

atgattgtga aagcaaaatt cgtgaaagga ttcattcgcg atgtgcaccc ttatgggtgc 60

cgccgtgaag ttctgaatca gatcgactac tgcaaaaaag ccattggctt tcgcggecca 120

aagaaagtgc tgatcggttg tgcttcctct ggcttcggtc tggctaccgc catttcctgt 180

gcggtcgggt gccacagaag ccacactatc ggcgctcagct atgaaaccgg tgcgaccgat 240

cgccgtattg gcacagcagg gtggtataac aatattttct ttaaagaatt tgccaaaaag 300

aaaggcctgg tggcaaaaaa ctttatcgaa gacgccttct cgaacgaaac caaggacaaa 360

gtcatcaaat atattaaaga cgaatttggc aaaatcgatc tgttcgttta ctgctggca 420

gcaccgcgtc gtaaggatta taagactggg aacgtttata cctcacgtat taaaacgatc 480

ctgggtgatt ttgaagggcc gactatcgat gtggaacgtg atgaaattac actgaaaaag 540

gtctcatctg cgtcaatcga agagattgaa gaaaccgta aggtgatggg cggcggaagat 600

tggcaagagt ggtgtgaaga actgctgtac gaagattgtt tcagtataa agccaccacc 660

atgccttatt cctatatcgg ttctcctcgc acctacaaaa tctaccgcga aggcactatc 720

ggcattgcga aaaaggatct ggaagataag gcaaaactga tcaacgagaa gctgaatcgc 780

gtcattggcg ggcgcgccatt cgttagcgtg aataaagccc tggttactaa ggcgagcgca 840

tatattccga ccttctcctc gtacgcccga attctgtata aagttatgaa agaaaagaat 900

attcacgaaa actgcattat gcaaattgaa cgcattgttt ccgagaaaaa ttattcaaat 960

gaaaagattc aatttgatga taaaggtcgt ctgcgtatgg atgacctgga gctgcgtaag 1020

gatgttcagg atgaagtaga ccgtatttgg agcaatatta caccggagaa ttttaaggaa 1080

ctgagcgact ataaaggcta caaaaaagaa tttatgaacc tgaatggatt tgatctggac 1140

ggcggtgatt attcaaagga tctggacatt gaactgctgc gcaaaactgga accataa 1197

<210> SEQ ID NO 77

<211> LENGTH: 1224

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Condon optimized EgTER

-continued

<400> SEQUENCE: 77

atggcgatgt ttacgaccac cgcaaaagtt attcagccga aaattcgtgg ttttatttgc	60
accaccacccc acccgattgg ttgcgaaaaa cgtgttcagg aagaaatcgc atacgcacgc	120
gcgcacccgc cgaccagccc gggtcgaaa cgtgtgctgg ttattggctg cagtacgggc	180
tatggcctga gcacccgtat caccgcggcc ttggttatac aggcgcgaac cctgggcgtg	240
tttctggcag gcccgccgac caaaggccgt cgggcgcgg cggttggtta taatacgggt	300
gcgttcgaaa aagccgcctt ggaagcaggt ctgtatgcac gttctctgaa tggatgatgcg	360
ttcgattcta ccacgaaagc ccgcaccgtg gaagcaatta aacgtgatct gggtagcgtt	420
gatctgggtg tgtatagcat tgcagcgccg aaacgtaccg atccggccac cggcgtgctg	480
cataaagcgt gcctgaaacc gattggtgca acctacacca atcgtacggg gaacaccgat	540
aaagcagaag ttaccgatgt gagtattgaa ccggccagtc cggaagaaat cgcagatacc	600
gtgaaagtta tgggtggcga agattgggaa ctgtggattc aggcactgag cgaagccggc	660
gtgctggcgc aaggcgcaaa aaccgttgcg tattcttata ttggcccgga aatgacgtgg	720
ccggtgtatt ggagtggcac cattggcgaa gccaaaaaag atgttgaaaa agcggcgaaa	780
cgcacacccc agcagtagcg ctgtccggcg tatccggttg ttgccaaagc gctggtgacc	840
caggccagta gcgccattcc ggtgggtgcg ctgtatatat gcctgctgta tcgtgttatg	900
aaagaaaaag gcacccatga aggctgcatt gaacagatgg tgcgtctgct gacgacgaaa	960
ctgtatccgg aaaatggtgc gccgatcgtg gatgaagcgg gccgtgtgcg tgttgatgat	1020
tgggaaatgg cagaagatgt tcagcaggca gttaaagatc tgtggagcca ggtgagtacg	1080
gccaatctga aagatattag cgattttgca ggttatcaga ccgaatttct gcgtctgttt	1140
ggctttggta ttgatggtgt ggattacgat cagccgggtg atgttgaagc ggatctgccg	1200
agcgcgcgcc agcagtaagt cgac	1224

<210> SEQ ID NO 78

<211> LENGTH: 1498

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Codon optimized ald gene

<400> SEQUENCE: 78

cggtagcctcg cgaatgcata tagatccaat catgcccggg ggtagcacia ggaggaaatag	60
ttcatgaata aagatacgtt gattccgacc acgaaagatc tgaaagtga aaccaacggt	120
gaaaaatatca acctgaaaaa ttataaagat aatagcagct gcttcggcgt gtttgaaaat	180
gttgaaaacg ccattttctc tgcggttcac gcacagaaaa tcctgtctct gcactatacc	240
aaagaacagc gcgaaaaaat cattaccgaa attcgcaaa cggccctgca gaataaagaa	300
gttctggcga ccatgatcct ggaagaaacc catatgggtc gttacgaaga taaaatcctg	360
aaacatgaac tgggtggcga atacacgcgc ggtacggaag atctgaccac gaccgcatgg	420
agcggtgata acggtctgac cgtgggtgaa atgtctccgt atggcgatat tggcgcaatt	480
acgccgagca cgaatccgac cgaacgggtt atttgtaaca gtattggcat gattgcagcc	540
ggtaaatgag tggttttcaa tggtagccg tgcgcacaaa aatgtgtggc gtttgccgtt	600
gaaatgatta acaaaagcgt tattagctgc ggtggccggg aaaatctggg gacgacgatt	660
aaaaaccgca ccatggaag tctggatgcg attattaaac atccgtctat taaactgctg	720
tgtggtaccg gcggtccggg tatggtgaaa accctgctga attctggcaa aaaagcaatt	780

-continued

```

ggcgcggtg cggtgaatcc gccggttatc gttgatgata ccgcggtatg tgaagaagca 840
ggccgcagta ttattgaagg ttgtagcttt gataataacc tgccgtgcat tgccgaaaaa 900
gaagtgtttg ttttcgaaaa tgtggcggat gatctgatca gcaacatgct gaaaaataac 960
gccgtgatta ttaacgaaga tcagggtgtc aaactgattg atctggttct gcagaaaaac 1020
aacgaaacgc aggaatatct cattaataaa aaatgggttg gtaaagatgc gaaactgttc 1080
ctggatgaaa tcgatgtgga aagtccgagc aacgtgaaat gtatcatctg cgaagtgaat 1140
gcccaaccatc cgtttgttat gacggaactg atgatgccga ttctgccgat tgttcgtgtt 1200
aaagatatgt atgaagccat caaatatgag aaaattgcgg aacagaaccg caaacacagc 1260
gcatatatct acagtaaaaa catcgataac ctgaaccgtt tcgaacgtga aattgatacc 1320
accatctttg tgaagaatgc caaaagtctt gcgggtgttg gctatgaagc cgaaggtctt 1380
accacgttca ccattgcagg ttctaccggc gaaggtatta ccagcgcgcg taattttacc 1440
cgccagcgcc gctgtgtgct ggcggtttaa gttaacccaa tatcggtacc cgggcccc 1498

```

```

<210> SEQ ID NO 79
<211> LENGTH: 6509
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: plasmid pPP988

```

```

<400> SEQUENCE: 79

```

```

tcgagggccc gcacatacga aaagactggc tgaagaacatt gagcctttga tgactgatga 60
tttggtgtaa gaagtggatc gattgtttga gaaaagaaga agaccataaa aataccttgt 120
ctgtcatcag acaggggtatt ttttatgctg tccagactgt ccgctgtgta aaaaatagga 180
ataaaggggg gttgttatta ttttactgat atgtaaaata taatttgtat aaggaattgt 240
gagcggtata caattcctac gaaaatgaga gggagaggaa acatgattca aaaacgaaag 300
cggacagttt cgttcagact tgtgcttatg tgcacgctgt tatttgtcag ttgcccatt 360
acaaaaacat cagccggtac ccaccatcac catcaccatt aagaattcct agaaactcca 420
agctatcttt aaaaaatcta gtaaatgcac gagcaacatc tttgtgtgct cagtgcattt 480
tttattttgt aactagata tttcttctcc gcttaaatca tcaaagaaat ctttatcact 540
tgtaaccagt ccgtccacat gtcgaattgc atctgaccga attttacgtt tccctgaata 600
attctcatca atcgtttcat caattttatc tttatacttt atattttgtg cgttaatcaa 660
atcataattt ttatatgttt cctcatgatt tatgtcttta ttattatagt tttattctc 720
tctttgatta tgtctttgta tcccgtttgt attacttgat cctttaactc tggcaaccct 780
caaaattgaa tgagacatgc tacacctccg gataataaat atatataaac gtatatagat 840
ttcataaagt ctaacacact agacttattt acttcgtaat taagtcgtta aaccgtgtgc 900
tctacgacca aaactataaa acctttaaga actttctttt tttacaagaa aaaagaaatt 960
agataaatct ctcatatctt ttattcaata atcgcatccg attgcagtat aaatttaacg 1020
atcactcatc atgttcatat ttatcagagc tcgtgctata attatactaa ttttataagg 1080
aggaaaaaat atgggcatct ttagtatttt tgtaatcagc acagttcatt atcaacaaaa 1140
caaaaaataa gtggttataa tgaatcgta ataagcaaaa ttcataaac caaattaaag 1200
agggttataa tgaacgagaa aaatataaaa cacagtcaaa actttattac ttcaaacat 1260
aatatagata aaataatgac aaatataaga ttaaatgaac atgataatat ctttgaaatc 1320

```

-continued

ggctcaggaa aaggccattt tacccttgaa ttagtaaaga ggtgtaattt cgtaactgcc	1380
attgaaatag accataaatt atgcaaaact acagaaaata aacttggtga tcacgataat	1440
ttccaagttt taaacaagga tatattgcag tttaaatctt ctaaaaacca atcctataaa	1500
atatatggta atatacctta taacataagt acggatataa tacgcaaaat tgtttttgat	1560
agtatagcta atgagattta tttaatcgtg gaatacgggt ttgctaaaag attattaaat	1620
acaaaacgct cattggcatt acttttaagt gcagaagttg atatttctat attaagtatg	1680
gttccaagag aatattttca tcttaaacct aaagtgaata gtcacttat cagattaagt	1740
agaaaaaat caagaatatt acacaagat aaacaaaagt ataattattt cgttatgaaa	1800
tgggttaaca aagaatacaa gaaaatattt acaaaaaatc aatttaacaa ttccttaaaa	1860
catgcaggaa ttgacgattt aaacaatatt agctttgaac aattcttatt tcttttcaat	1920
agctataaat tatttaataa gtaagttaag ggatgcagtt catcgatgaa ggcaactaca	1980
gctcaggcga caaccatacg ctgagagatc ctcactacgt agaagataaa ggccacaaat	2040
acttagtatt tgaagcaaac actggaactg aagatggcta ccaaggcgaa gaatctttat	2100
ttaacaaagc atactatggc aaaagcacat cattcttcgg tcaagaaagt caaaaacttc	2160
tgcaaacgga taaaaaacgc acggctgagt tagcaaacgg cgctctcggg atgattgagc	2220
taaacgatga ttacacactg aaaaaagtga tgaaaccgct gattgcatct aacacagtaa	2280
cagatgaaat tgaacgcgcg aacgtcttta aaatgaacgg caaatggtag ctgttcactg	2340
actcccgcgg atcaaaaatg acgattgacg gcattacgtc taacgatatt tacatgcttg	2400
gttatgtttc taattcttta actggcccat acaagccgct gaacaaaact ggccctgtgt	2460
taaaaatgga tcttgatcct aacgatgtaa cctttactta ctcacacttc gctgtacctc	2520
aagcgaaagg aaacaatgac gtgattacaa gctatatgac aaacagagga ttctacgcag	2580
acaaacaatc aacgtttgcg ccaagcttgc atgcgagagt agggaaactgc caggcatcaa	2640
ataaacgaa aggcctcagc gaaagactgg gcctttcgtt ttatctgttg ttgtcggtg	2700
aacgtctccc tgagtaggac aaatccgcgg ggagcggatt tgaacgttgc gaagcaacgg	2760
cccgagggtt ggcgggcagg acgcccgcga taaactgccg ggcatcaaat taagcagaag	2820
gccatcctga cggatggcct ttttgcgttt ctacaaactc tttttgttta tttttctaaa	2880
tacattcaaa tatgtatccg ctcatgctcc ggatctgcat cgcaggatgc tgetggctac	2940
cctgtggaac acctacatct gtattaacga agcgctggca ttgaccctga gtgatttttc	3000
tctggtcccg ccgcatccat accgccagtt gtttaccctc acaacgttcc agtaaccggg	3060
catgttcacg atcagtaacc cgtatcgtga gcatacctc tcgtttcacg ggtatcatta	3120
cccccatgaa cagaaattcc cccttacacg gaggcacaa gtgaccaaag aggaaaaaac	3180
cgcccttaac atggcccgcg ttatcagaag ccagacatta acgcttctgg agaaactcaa	3240
cgagctggac gcggatgaac aggcagacat ctgtgaatcg cttcacgacc acgctgatga	3300
gctttaccgc agctgcctcg cgcgtttcgg tgatgacggg gaaaacctct gacacatgca	3360
gctcccggag acggtcacag cttgtctgta agcggatgcc gggagcagac aagcccgtca	3420
gggcgcgtca gcgggtgttg gcgggtgtcg gggcgcagcc atgaccagc cacgtagcga	3480
tagcggagtg tatactggct taactatcg gcacagagc agattgtact gagagtgcac	3540
catatgcggg gtgaaatacc gcacagatgc gtaaggagaa aataccgcac caggcgctct	3600
tccgcttccc cgctcactga ctgcgtgcgc tcggctcgtc ggctgcccgc agcggatatca	3660
gctcactcaa aggcggtaat acggttatcc acagaatcag gggataacgc aggaaagaac	3720

-continued

atgtgagcaa aagccagca aaaggccagg aaccgtaaaa aggcgcggtt gctggcggtt	3780
ttccataggc tccgcccccc tgacgagcat cacaaaaatc gacgctcaag tcagagggtg	3840
cgaaacccga caggactata aagataccag gcgtttcccc ctggaagctc cctcgtgcgc	3900
tctcctgttc cgaccctgcc gcttacggga tacctgtccg cctttctccc ttcgggaagc	3960
gtggcgcttt ctcaatgctc acgctgtagg tatctcagtt cgggtgtaggt cgttcgtcc	4020
aagctgggct gtgtgcacga acccccgtt cagcccgacc gctgcgcctt atccggtaac	4080
tatcgtcttg agtccaaccc ggtaagacac gacttatcgc cactggcagc agccactggt	4140
aacaggatta gcagagcgag gtatgtaggc ggtgctacag agttcttgaa gtggtggcct	4200
aactacggct aactagaag gacagtattt ggtatctgcg ctctgctgaa gccagttacc	4260
ttcggaaaaa gagttgtag ctcttgatcc ggcaaaaaa ccaccgctgg tagcgggtgt	4320
ttttttgttt gcaagcagca gattacgcgc agaaaaaaag gatctcaaga agatcctttg	4380
atcttttcta cgggggtctga cgctcagtggt aacgaaaact caggttaagg gattttgttc	4440
atgagattat caaaaaggat cttcacctag atccttttaa attaaaaatg aagttttaaa	4500
tcaatctaaa gtatatatga gtaaaacttg tctgacagtt accaatgctt aatcagttag	4560
gcacatatct cagcgatctg tctatttcgt tcatccatag ttgcctgact ccccgctgtg	4620
tagataacta cgatacggga gggcctacca tctggcccca gtgctgcaat gataccgcga	4680
gaccacgct caccggctcc agatttatca gcaataaacc agccagccgg aagggccgag	4740
cgcagaagtg gtcttgcaac tttatccgcc tccatccagt ctattaattg ttgccgggaa	4800
gctagagtaa gtatgtgcc agttaatagt ttgcgcaacg ttgttgccat tgctgcaggc	4860
atcgtggtgt cacgctcgtc gtttggtatg gcttcattca gctccggttc ccaacgatca	4920
aggcgagtta catgatcccc catgttgtgc aaaaaagcgg ttagctcctt cggctcctccg	4980
atcgttgta gaagtaagtt ggccgcagtg ttatcactca tggttatggc agcactgcat	5040
aattctctta ctgtcatgcc atccgtaaga tgcttttctg tgactggtga gtactcaacc	5100
aagtcattct gagaatagtg tatgcggcga ccgagttgct cttgcccggc gtcaatacgg	5160
gataataccg cgccacatag cagaacttta aaagtgtcga tcattggaaa acgttcttcg	5220
gggcgaaaac tctcaaggat cttaccgctg ttgagatcca gttcgatgta acccactcgt	5280
gcacccaact gatcttcagc atcttttact ttcaccagcg tttctgggtg agcaaaaaca	5340
ggaaggcaaa atgccgcaaa aaagggaata agggcgacac ggaaatgttg aatactcata	5400
ctcttccttt ttcaatatta tgaagcatt tatcagggtt attgtctcat gagcggatac	5460
atatttgaat gtatttagaa aaataaaca ataggggttc cgcgcacatt tccccgaaaa	5520
gtgccacctg acgtctaaga aaccattatt atcatgacat taacctataa aaataggcgt	5580
atcacgaggc cctttcgtct cgcgcgtttc ggtgatgacg gtgaaaacct ctgacacatg	5640
cagctccccg agacggctcag agcttgtctg taagcggatg ccgggagcag acaagcccgt	5700
cagggcgcgt cagcgggtgt tcatgtgcgt aactaacttg ccatcttcaa acaggagggc	5760
tggaagaagc agaccgctaa cacagtacat aaaaaaggag acatgaacga tgaacatcaa	5820
aaagtttgca aaacaagcaa cagtattaac ctttactacc gcaactgctgg caggaggcgc	5880
aactcaagcg tttgcgaaag aaacgaacca aaagccatat aaggaaacat acggcatttc	5940
ccatattaca cgccatgata tgctgcaaat ccttgaacag caaaaaaatg aaaaatatca	6000
agttcctgaa ttcgattcgt ccacaattaa aaatatctct tctgcaaaag gcctggacgt	6060

-continued

```

ttgggacagc tggccattac aaaacgctga cggcactgtc gcaaactatc acggctacca 6120
catcgtcttt gcattagccg gagatcctaa aaatgcggat gacacatcga tttacatgtt 6180
ctatcaaaaa gtcggcgaaa cttctattga cagctggaaa aacgctggcc gcgtctttaa 6240
agacagcgac aaattcgatg caaatgattc taccctaaaa gaccaaacac aagaatggtc 6300
agggttcagcc acatttcatc ctgacggaaa aatccgttta ttctacactg atttctccgg 6360
taaacattac ggcaacaaaa cactgacaac tgcacaagtt aacgtatcag catcagacag 6420
ctctttgaac atcaacgggtg tagaggatta taaatcaatc tttgacgggtg acggaaaaac 6480
gtatcaaaat gtacagcatg ccacgcgtc 6509

```

```

<210> SEQ ID NO 80
<211> LENGTH: 46
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: primer N85

```

```

<400> SEQUENCE: 80

```

```

catagatctg gatccaaagg aggggtgagga aatggcgatg tttacg 46

```

```

<210> SEQ ID NO 81
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: primer N86

```

```

<400> SEQUENCE: 81

```

```

gtcgcacttac tgctgggcgg 20

```

```

<210> SEQ ID NO 82
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: T7 primer

```

```

<400> SEQUENCE: 82

```

```

taatacgact cactataggg 20

```

```

<210> SEQ ID NO 83
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer Trc99af

```

```

<400> SEQUENCE: 83

```

```

ttgacaatta atcatccggc 20

```

```

<210> SEQ ID NO 84
<211> LENGTH: 21
<212> TYPE: DNA
<213> ORGANISM: artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer N5SeqF4

```

```

<400> SEQUENCE: 84

```

```

ggtcactgt tccggaaatt c 21

```

```

<210> SEQ ID NO 85
<211> LENGTH: 46

```

-continued

<212> TYPE: DNA
<213> ORGANISM: artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer T-ald(BamHI)

<400> SEQUENCE: 85
tgatctggat ccaagaagga gcccttcacc atgaataaag acacac 46

<210> SEQ ID NO 86
<211> LENGTH: 54
<212> TYPE: DNA
<213> ORGANISM: artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer B-ald(ETGER)

<400> SEQUENCE: 86
catcgccatt tctcaccct cctttttagc cggcaagtac acatcttctt tgtc 54

<210> SEQ ID NO 87
<211> LENGTH: 42
<212> TYPE: DNA
<213> ORGANISM: artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer T-Ptrc(BspEI)

<400> SEQUENCE: 87
ttccgtactt ccggacgact gcacggtgca ccaatgcttc tg 42

<210> SEQ ID NO 88
<211> LENGTH: 43
<212> TYPE: DNA
<213> ORGANISM: artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer B-aldopt(BspEI)

<400> SEQUENCE: 88
cggatcttaa gtactttaac ccgccagcac acagcggcgc tgg 43

<210> SEQ ID NO 89
<211> LENGTH: 32
<212> TYPE: DNA
<213> ORGANISM: artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer T-BspEIAatII

<400> SEQUENCE: 89
ccggatcatg ataataatgg tttcttagac gt 32

<210> SEQ ID NO 90
<211> LENGTH: 24
<212> TYPE: DNA
<213> ORGANISM: artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer B-BspEIAatII

<400> SEQUENCE: 90
ctaagaaacc attattatca tgat 24

<210> SEQ ID NO 91
<211> LENGTH: 42
<212> TYPE: DNA
<213> ORGANISM: artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: 1.6GI promoter

<400> SEQUENCE: 91

-continued

gcccttgaca atgccacatc ctgagcaaat aattcaacca ct 42

<210> SEQ ID NO 92
 <211> LENGTH: 42
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: 1.5GI promoter

<400> SEQUENCE: 92

gcccttgact atgccacatc ctgagcaaat aattcaacca ct 42

<210> SEQ ID NO 93
 <211> LENGTH: 39
 <212> TYPE: DNA
 <213> ORGANISM: artificial sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Primer AFBamHI

<400> SEQUENCE: 93

cattggatcc atgaataaag acacactaat acctacaac 39

<210> SEQ ID NO 94
 <211> LENGTH: 39
 <212> TYPE: DNA
 <213> ORGANISM: artificial sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Primer ARAat2

<400> SEQUENCE: 94

catgacgtca ctagtggttaa caagaagtta gccggaag 39

<210> SEQ ID NO 95
 <211> LENGTH: 50
 <212> TYPE: DNA
 <213> ORGANISM: artificial sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Primer Forward 1(E)

<400> SEQUENCE: 95

catgttaaca aaggaggaaa gatctatggc gatgtttacg accaccgcaa 50

<210> SEQ ID NO 96
 <211> LENGTH: 43
 <212> TYPE: DNA
 <213> ORGANISM: artificial sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Primer bottom reverse 1(E)

<400> SEQUENCE: 96

cccctccttt ggcgcgcctt actgctgggc ggcgctcggc aga 43

<210> SEQ ID NO 97
 <211> LENGTH: 51
 <212> TYPE: DNA
 <213> ORGANISM: artificial sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Primer Top Forward 2(B)

<400> SEQUENCE: 97

gcccgagcagt aaggcgcgcc aaaggagggg ttaaaatggt tgatttcgaa t 51

<210> SEQ ID NO 98
 <211> LENGTH: 42
 <212> TYPE: DNA
 <213> ORGANISM: artificial sequence

-continued

<220> FEATURE:
 <223> OTHER INFORMATION: Primer reverse 2(B)

 <400> SEQUENCE: 98

 gtcgacgtca tactagttta cacagatttt ttgaatattt gt 42

 <210> SEQ ID NO 99
 <211> LENGTH: 47
 <212> TYPE: DNA
 <213> ORGANISM: artificial sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Primer Pamy/lacOF

 <400> SEQUENCE: 99

 cattgtacag aattcgagct ctcgaggccc cgcacatacg aaaagac 47

 <210> SEQ ID NO 100
 <211> LENGTH: 52
 <212> TYPE: DNA
 <213> ORGANISM: Artificial sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Primer Pam/la cOR

 <400> SEQUENCE: 100

 cattgtacag tttaaacata ggtcaccctc attttcgtag gaattgttat cc 52

 <210> SEQ ID NO 101
 <211> LENGTH: 52
 <212> TYPE: DNA
 <213> ORGANISM: artificial sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Primer Spac F

 <400> SEQUENCE: 101

 cattgtacag tttaaacata ggtcaccctc attttcgtag gaattgttat cc 52

 <210> SEQ ID NO 102
 <211> LENGTH: 36
 <212> TYPE: DNA
 <213> ORGANISM: artificial sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Primer Spac R

 <400> SEQUENCE: 102

 catgtttaaa cggtgaccca agctggggat cgcgcg 36

 <210> SEQ ID NO 103
 <211> LENGTH: 44
 <212> TYPE: DNA
 <213> ORGANISM: artificial sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Primer Top TF

 <400> SEQUENCE: 103

 cattggtcac cattcccgagg catgcaaagg aggttagtag aatg 44

 <210> SEQ ID NO 104
 <211> LENGTH: 51
 <212> TYPE: DNA
 <213> ORGANISM: artificial sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Primer Bot TR

 <400> SEQUENCE: 104

 cctttacgag accggtacta gtcaagtcga cagggcgcgc ccaatacttt c 51

-continued

<210> SEQ ID NO 105
<211> LENGTH: 57
<212> TYPE: DNA
<213> ORGANISM: artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer Top CF

<400> SEQUENCE: 105

cgcgccctgt cgacttgact agtaccggtc gcgtaaagga ggtattagtc atggaac 57

<210> SEQ ID NO 106
<211> LENGTH: 38
<212> TYPE: DNA
<213> ORGANISM: artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer Bot CR

<400> SEQUENCE: 106

catcgtttaa acttgatcc agatccctta cctcctat 38

<210> SEQ ID NO 107
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer N3SeqF1

<400> SEQUENCE: 107

ccatcatacc atactgaccc 20

<210> SEQ ID NO 108
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer N3SeqF2

<400> SEQUENCE: 108

gctactggag cattgctcac 20

<210> SEQ ID NO 109
<211> LENGTH: 24
<212> TYPE: DNA
<213> ORGANISM: artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer N3SeqF3

<400> SEQUENCE: 109

ccattaacag ctgctattac aggc 24

<210> SEQ ID NO 110
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer N4SeqR3

<400> SEQUENCE: 110

ggctctcgaa taacacctgg 20

<210> SEQ ID NO 111
<211> LENGTH: 22
<212> TYPE: DNA
<213> ORGANISM: artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer N5SeqF3

-continued

<400> SEQUENCE: 111

caagcttcat aacaggagct gg

22

<210> SEQ ID NO 112

<211> LENGTH: 22

<212> TYPE: DNA

<213> ORGANISM: artificial sequence

<220> FEATURE:

<223> OTHER INFORMATION: Primer N7SeqR2

<400> SEQUENCE: 112

atcccacaat cgcgcagtga tc

22

<210> SEQ ID NO 113

<211> LENGTH: 20

<212> TYPE: DNA

<213> ORGANISM: artificial sequence

<220> FEATURE:

<223> OTHER INFORMATION: Primer N31SeqF1

<400> SEQUENCE: 113

ctgagataag aaagcccgca

20

<210> SEQ ID NO 114

<211> LENGTH: 20

<212> TYPE: DNA

<213> ORGANISM: artificial sequence

<220> FEATURE:

<223> OTHER INFORMATION: Primer N62SeqF2

<400> SEQUENCE: 114

caaccctggg cgtgtttctg

20

<210> SEQ ID NO 115

<211> LENGTH: 20

<212> TYPE: DNA

<213> ORGANISM: artificial sequence

<220> FEATURE:

<223> OTHER INFORMATION: Primer N62SeqF3

<400> SEQUENCE: 115

gtggcgaaga ttgggaactg

20

<210> SEQ ID NO 116

<211> LENGTH: 24

<212> TYPE: DNA

<213> ORGANISM: artificial sequence

<220> FEATURE:

<223> OTHER INFORMATION: Primer N62SeqF4

<400> SEQUENCE: 116

gggaaatggc agaagatggt cagc

24

<210> SEQ ID NO 117

<211> LENGTH: 24

<212> TYPE: DNA

<213> ORGANISM: artificial sequence

<220> FEATURE:

<223> OTHER INFORMATION: Primer N63SeqR1

<400> SEQUENCE: 117

cggctctgata acctgcaaaa tcgc

24

<210> SEQ ID NO 118

-continued

<211> LENGTH: 21
<212> TYPE: DNA
<213> ORGANISM: artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer N63SeqR2

<400> SEQUENCE: 118

caccagcgct ttggcaacaa c 21

<210> SEQ ID NO 119
<211> LENGTH: 23
<212> TYPE: DNA
<213> ORGANISM: artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer N63SeqR3

<400> SEQUENCE: 119

gaacgtgcat acagacctgc ttc 23

<210> SEQ ID NO 120
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer N63SeqR4

<400> SEQUENCE: 120

cggctgaata acttttgagg 20

<210> SEQ ID NO 121
<211> LENGTH: 24
<212> TYPE: DNA
<213> ORGANISM: artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer Pamy SeqF2

<400> SEQUENCE: 121

gcctttgatg actgatgatt tggc 24

<210> SEQ ID NO 122
<211> LENGTH: 24
<212> TYPE: DNA
<213> ORGANISM: artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer Pamy SeqF

<400> SEQUENCE: 122

tctccggtaa acattacggc aaac 24

<210> SEQ ID NO 123
<211> LENGTH: 24
<212> TYPE: DNA
<213> ORGANISM: artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer Pamy seqR

<400> SEQUENCE: 123

cggtcagatg caattcgaca tgtg 24

<210> SEQ ID NO 124
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: artificial sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer SpacF Seq

<400> SEQUENCE: 124

-continued

gaagtgtgca agacctcact	20
<210> SEQ ID NO 125 <211> LENGTH: 24 <212> TYPE: DNA <213> ORGANISM: artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Primer sacB Up <400> SEQUENCE: 125	
cggttttgtt actgataaag cagg	24
<210> SEQ ID NO 126 <211> LENGTH: 24 <212> TYPE: DNA <213> ORGANISM: artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Primer sacB Dn <400> SEQUENCE: 126	
cggttagcca ttgcctgct tttta	24
<210> SEQ ID NO 127 <211> LENGTH: 22 <212> TYPE: DNA <213> ORGANISM: artificial sequence <220> FEATURE: <223> OTHER INFORMATION: Primer HT R <400> SEQUENCE: 127	
acaaagatct ccatggaagc gt	22
<210> SEQ ID NO 128 <211> LENGTH: 1185 <212> TYPE: DNA <213> ORGANISM: Escherichia coli <400> SEQUENCE: 128	
atgaaaaatt gtgtcatcgt cagtgcggta cgtactgcta tcggtagttt taacggttca	60
ctcgtttcca ccagcgccat cgacctgggg gcgacagtaa ttaaagccgc cattgaacgt	120
gcaaaaatcg attcacaaca cgttgatgaa gtgattatgg gtaacgtgtt acaagccggg	180
ctggggcaaa atccggcgcg tcaggcactg ttaaaaagcg ggctggcaga aacgggtgtc	240
ggattcacgg tcaataaagt atgtggttcg ggtcttaaaa gtgtggcgct tgccgccag	300
gccattcagg caggtcaggc gcagagcatt gtggcggggg gtatggaaaa tatgagttta	360
gccccctact tactcgatgc aaaagcacgc tctggttacc gtcttgaga cggacagggt	420
tatgacgtaa tctcgcgga tggcctgatg tgcgccaccc atggttatca tatggggatt	480
accgccgaaa acgtggctaa agagtacgga attacccgtg aaatgcagga tgaactggcg	540
ctacattcac agcgtaaagc ggcagccgca attgagtcgg gtgtttttac agccgaaatc	600
gtcccggtaa atgttgtcac tcgaaagaaa accttcgtct tcagtcaaga cgaattcccg	660
aaagcgaatt caacggctga agcgttaggt gcattgcgcc cggccttcga taaagcagga	720
acagtcaccg ctgggaacgc gtctggtatt aacgacggtg ctgccgctct ggtgattatg	780
gaagaatctg cggcgctggc agcaggcctt acccccctgg ctgcgattaa aagttatgcc	840
agcggtgggc tgccccccgc attgatgggt atggggccag tacctgccac gcaaaaagcg	900
ttacaactgg cggggctgca actggcggat attgatctca ttgaggctaa tgaagcattt	960

-continued

```

gctgcacagt tccttgccgt tgggaaaaac ctgggctttg attctgagaa agtgaatgtc 1020
aacggcgggg ccctcgcgct cgggcctcct atcggtgcca gtggtgctcg tattctggtc 1080
acactattac atgccatgca ggcaacgatg aaaacgctgg ggctggcaac actgtgcatt 1140
ggcgcggtc agggaattgc gatggtgatt gaacggttga attaa 1185

```

<210> SEQ ID NO 129

<211> LENGTH: 394

<212> TYPE: PRT

<213> ORGANISM: Escherichia coli

<400> SEQUENCE: 129

```

Met Lys Asn Cys Val Ile Val Ser Ala Val Arg Thr Ala Ile Gly Ser
1           5           10          15
Phe Asn Gly Ser Leu Ala Ser Thr Ser Ala Ile Asp Leu Gly Ala Thr
20          25          30
Val Ile Lys Ala Ala Ile Glu Arg Ala Lys Ile Asp Ser Gln His Val
35          40          45
Asp Glu Val Ile Met Gly Asn Val Leu Gln Ala Gly Leu Gly Gln Asn
50          55          60
Pro Ala Arg Gln Ala Leu Leu Lys Ser Gly Leu Ala Glu Thr Val Cys
65          70          75          80
Gly Phe Thr Val Asn Lys Val Cys Gly Ser Gly Leu Lys Ser Val Ala
85          90          95
Leu Ala Ala Gln Ala Ile Gln Ala Gly Gln Ala Gln Ser Ile Val Ala
100         105         110
Gly Gly Met Glu Asn Met Ser Leu Ala Pro Tyr Leu Leu Asp Ala Lys
115         120         125
Ala Arg Ser Gly Tyr Arg Leu Gly Asp Gly Gln Val Tyr Asp Val Ile
130         135         140
Leu Arg Asp Gly Leu Met Cys Ala Thr His Gly Tyr His Met Gly Ile
145         150         155         160
Thr Ala Glu Asn Val Ala Lys Glu Tyr Gly Ile Thr Arg Glu Met Gln
165         170         175
Asp Glu Leu Ala Leu His Ser Gln Arg Lys Ala Ala Ala Ala Ile Glu
180         185         190
Ser Gly Ala Phe Thr Ala Glu Ile Val Pro Val Asn Val Val Thr Arg
195         200         205
Lys Lys Thr Phe Val Phe Ser Gln Asp Glu Phe Pro Lys Ala Asn Ser
210         215         220
Thr Ala Glu Ala Leu Gly Ala Leu Arg Pro Ala Phe Asp Lys Ala Gly
225         230         235         240
Thr Val Thr Ala Gly Asn Ala Ser Gly Ile Asn Asp Gly Ala Ala Ala
245         250         255
Leu Val Ile Met Glu Glu Ser Ala Ala Leu Ala Ala Gly Leu Thr Pro
260         265         270
Leu Ala Arg Ile Lys Ser Tyr Ala Ser Gly Gly Val Pro Pro Ala Leu
275         280         285
Met Gly Met Gly Pro Val Pro Ala Thr Gln Lys Ala Leu Gln Leu Ala
290         295         300
Gly Leu Gln Leu Ala Asp Ile Asp Leu Ile Glu Ala Asn Glu Ala Phe
305         310         315         320
Ala Ala Gln Phe Leu Ala Val Gly Lys Asn Leu Gly Phe Asp Ser Glu
325         330         335

```

Lys Val Asn Val Asn Gly Gly Ala Ile Ala Leu Gly His Pro Ile Gly
340 345 350

Ala Ser Gly Ala Arg Ile Leu Val Thr Leu Leu His Ala Met Gln Ala
355 360 365

Arg Asp Lys Thr Leu Gly Leu Ala Thr Leu Cys Ile Gly Gly Gly Gln
370 375 380

Gly Ile Ala Met Val Ile Glu Arg Leu Asn
385 390

```
<210> SEQ ID NO 130
<211> LENGTH: 1182
<212> TYPE: DNA
<213> ORGANISM: Bacillus subtilis
```

<400> SEQUENCE: 130

ttaaataaac	tgcactaaga	cggcgtctcc	ctgtgtctgc	ccgctgcaaa	tagcggcaac	60
gcccgagccc	cctccccgct	gctttaattc	ataaacaagc	gtcatgagaa	ttctcgccc	120
gctcgcgccg	atcggttggc	cgagcgcgat	cgcaccgcc	ttcacattta	ctttttcaag	180
atcgaaaacct	acgatttttt	cacatgtcaa	aacaactgaa	gcaaaagctt	catttacttc	240
aaacaagtca	atatcttgga	cagttaaacc	attctttttc	aggagcttgt	taatagcaaa	300
ccctggcgct	gccgccagct	cgctgcgtgg	cattccccta	gttgaaaaac	caagaattgt	360
agccagaggc	cgtttgccaa	gctcagcagc	tttttctcca	gacatcagca	cgaacgcgcc	420
ggctccgtca	ttgactccag	gagcattgcc	ggctgtgata	gaaccgtcac	ttgcataaat	480
cggagcaagt	tttgcgagct	gatccagact	tgtgtcacgg	cgaatcgctt	catctttatc	540
aacaacgttt	ggttttcttt	ttcgaccgat	ccagttgacg	ggaacaattt	catcctgaaa	600
cttccttcca	tccggcgccct	tagctgccct	tgcattgact	ctcaacgccc	attcgtctct	660
ctctctctgt	gagattgcatt	attccttggc	agctgtattt	ccgtgaacac	tcacgtgcac	720
ctcgtcaaat	gcgcacgtta	atccgtcata	caccattaa	tccttaagct	cgcgcgtccc	780
catccgtgct	ccccagcgcc	cggcgggaa	ggcatacgg	atattgtcca	tgttttccat	840
cccccccgca	acaagtatgt	cgcgcatcgt	cgcgcgaatc	atttgatcac	ataaagtgc	900
agcgcgaagg	ccggaagcac	agactttatt	cagtgtttct	gacggcacac	tccaaggcat	960
tcccgccaga	cgggcagctt	gacgggaagg	tatctgccct	gagccggcct	ggacaacccat	1020
gcccatgacg	tttccttcta	catcatctcc	agagactcca	gcctgttgca	gcgcctcctt	1080
catcacaatg	cccccaagct	cagcagcttt	cacctcttcc	aaaactccgc	cgaatttgcc	1140
aatggagatt	cttgacgacg	ttacaatgac	tgttttctcc	at		1180

```
<210> SEQ ID NO 131
<211> LENGTH: 393
<212> TYPE: PRT
<213> ORGANISM: Bacillus subtilis
```

<400> SEQUENCE: 131

Met Arg Lys Thr Val Ile Val Ser Ala Ala Arg Thr Pro Phe Gly Lys
1 5 10 15

Phe Gly Gly Val Leu Lys Glu Val Lys Ala Ala Glu Leu Gly Gly Ile
20 25 30

Val Met Lys Glu Ala Leu Gln Gln Ala Gly Val Ser Gly Asp Asp Val
35 40 45

Glu Gly Asn Val Met Gly Met Val Val Gln Ala Gly Ser Gly Gln Ile
50 55 60

-continued

Pro Ser Arg Gln Ala Ala Arg Leu Ala Gly Met Pro Trp Ser Val Pro
 65 70 75 80
 Ser Glu Thr Leu Asn Lys Val Cys Ala Ser Gly Leu Arg Ala Val Thr
 85 90 95
 Leu Cys Asp Gln Met Ile Arg Ala Gln Asp Ala Asp Ile Leu Val Ala
 100 105 110
 Gly Gly Met Glu Ser Met Ser Asn Ile Pro Tyr Ala Val Pro Ala Gly
 115 120 125
 Arg Trp Gly Ala Arg Met Gly Asp Gly Glu Leu Arg Asp Leu Met Val
 130 135 140
 Tyr Asp Gly Leu Thr Cys Ala Phe Asp Glu Val His Met Ala Val His
 145 150 155 160
 Gly Asn Thr Ala Ala Lys Glu Tyr Ala Ile Ser Arg Arg Glu Gln Asp
 165 170 175
 Glu Trp Ala Leu Arg Ser His Ala Arg Ala Ala Lys Ala Ala Asp Glu
 180 185 190
 Gly Lys Phe Gln Asp Glu Ile Val Pro Val Asn Trp Ile Gly Arg Lys
 195 200 205
 Gly Lys Pro Asn Val Val Asp Lys Asp Glu Ala Ile Arg Arg Asp Thr
 210 215 220
 Ser Leu Asp Gln Leu Ala Lys Leu Ala Pro Ile Tyr Ala Ser Asp Gly
 225 230 235 240
 Ser Ile Thr Ala Gly Asn Ala Pro Gly Val Asn Asp Gly Ala Gly Ala
 245 250 255
 Phe Val Leu Met Ser Glu Glu Lys Ala Ala Glu Leu Gly Lys Arg Pro
 260 265 270
 Leu Ala Thr Ile Leu Gly Phe Ser Thr Thr Gly Met Pro Ala His Glu
 275 280 285
 Leu Ala Ala Ala Pro Gly Phe Ala Ile Asn Lys Leu Leu Lys Lys Asn
 290 295 300
 Gly Leu Thr Val Gln Asp Ile Asp Leu Phe Glu Val Asn Glu Ala Phe
 305 310 315 320
 Ala Ser Val Val Leu Thr Cys Glu Lys Ile Val Gly Phe Asp Leu Glu
 325 330 335
 Lys Val Asn Val Asn Gly Gly Ala Ile Ala Leu Gly His Pro Ile Gly
 340 345 350
 Ala Ser Gly Ala Arg Ile Leu Met Thr Leu Val Tyr Glu Leu Lys Arg
 355 360 365
 Arg Gly Gly Gly Leu Gly Val Ala Ala Ile Cys Ser Gly Ala Ala Gln
 370 375 380
 Gly Asp Ala Val Leu Val Gln Val His
 385 390

<210> SEQ ID NO 132

<211> LENGTH: 1197

<212> TYPE: DNA

<213> ORGANISM: *Saccharomyces cerevisiae*

<400> SEQUENCE: 132

atgtctcaga acgtttacat tgtatcgact gccagaaccc caattgggttc attccagggt	60
tctctatcct ccaagacagc agtgaattg ggtgctgttg ctttaaaagg cgccttggt	120
aaggttccag aattggatgc atccaaggat ttgacgaaa ttatttttgg taacgttctt	180
tctgcccaatt tgggcccaagc tccggccaga caagttgctt tggctgccgg tttagtaat	240

-continued

```

catatcggtg caagcacagt taacaaggtc tgtgcatccg ctatgaaggc aatcattttg   300
ggtgctcaat ccatcaaatg tggtaatgct gatgtgtgctg tagctggtgg ttgtgaatct   360
atgactaacg caccatacta catgccagca gcccggtgcgg gtgccaaatt tggccaaact   420
gttcttgttg atggtgtcga aagagatggg ttgaacgatg cgtacgatgg tctagccatg   480
ggtgtacacg cagaaaagtg tgcccgtgat tgggatatta ctagagaaca acaagacaat   540
tttgccatcg aatcctacca aaaatctcaa aaatctcaaa aggaaggtaa attcgacaat   600
gaaattgtac ctgttaccat taagggattt agaggtgaag ctgatactca agtcacgaag   660
gacgaggaac ctgctagatt acacgttgaa aaattgagat ctgcaaggac tgttttccaa   720
aaagaaaacg gtactgttac tgccgctaac gcttctccaa tcaacgatgg tgctgcagcc   780
gtcatcttgg ttccgaaaaa agttttgaag gaaaagaatt tgaagccttt ggctattatc   840
aaaggttggg gtgaggccgc tcataacca gctgatttta catgggctcc atctcttgca   900
gttccaaagg ctttgaaaaca tgctggcatc gaagacatca attctgttga ttactttgaa   960
ttcaatgaag ccttttcggt tgctggtttg gtgaacacta agattttgaa gctagaccca  1020
tctaaggtta atgtatatgg tgggtgctgtt gctctaggtc acccattggg ttgttctggt  1080
gctagagtgg ttgttacact gctatccatc ttacagcaag aaggaggtaa gatcggtgtt  1140
gccgccattt gtaatggtgg tggtggtgct tcctctattg tcattgaaaa gatatga   1197

```

<210> SEQ ID NO 133

<211> LENGTH: 398

<212> TYPE: PRT

<213> ORGANISM: *Saccharomyces cerevisiae*

<400> SEQUENCE: 133

```

Met Ser Gln Asn Val Tyr Ile Val Ser Thr Ala Arg Thr Pro Ile Gly
1           5              10              15

Ser Phe Gln Gly Ser Leu Ser Ser Lys Thr Ala Val Glu Leu Gly Ala
20          25              30

Val Ala Leu Lys Gly Ala Leu Ala Lys Val Pro Glu Leu Asp Ala Ser
35          40              45

Lys Asp Phe Asp Glu Ile Ile Phe Gly Asn Val Leu Ser Ala Asn Leu
50          55              60

Gly Gln Ala Pro Ala Arg Gln Val Ala Leu Ala Ala Gly Leu Ser Asn
65          70              75              80

His Ile Val Ala Ser Thr Val Asn Lys Val Cys Ala Ser Ala Met Lys
85          90              95

Ala Ile Ile Leu Gly Ala Gln Ser Ile Lys Cys Gly Asn Ala Asp Val
100         105              110

Val Val Ala Gly Gly Cys Glu Ser Met Thr Asn Ala Pro Tyr Tyr Met
115         120              125

Pro Ala Ala Arg Ala Gly Ala Lys Phe Gly Gln Thr Val Leu Val Asp
130         135              140

Gly Val Glu Arg Asp Gly Leu Asn Asp Ala Tyr Asp Gly Leu Ala Met
145         150              155              160

Gly Val His Ala Glu Lys Cys Ala Arg Asp Trp Asp Ile Thr Arg Glu
165         170              175

Gln Gln Asp Asn Phe Ala Ile Glu Ser Tyr Gln Lys Ser Gln Lys Ser
180         185              190

Gln Lys Glu Gly Lys Phe Asp Asn Glu Ile Val Pro Val Thr Ile Lys
195         200              205

```

-continued

Gly Phe Arg Gly Lys Pro Asp Thr Gln Val Thr Lys Asp Glu Glu Pro
 210 215 220
 Ala Arg Leu His Val Glu Lys Leu Arg Ser Ala Arg Thr Val Phe Gln
 225 230 235 240
 Lys Glu Asn Gly Thr Val Thr Ala Ala Asn Ala Ser Pro Ile Asn Asp
 245 250 255
 Gly Ala Ala Ala Val Ile Leu Val Ser Glu Lys Val Leu Lys Glu Lys
 260 265 270
 Asn Leu Lys Pro Leu Ala Ile Ile Lys Gly Trp Gly Glu Ala Ala His
 275 280 285
 Gln Pro Ala Asp Phe Thr Trp Ala Pro Ser Leu Ala Val Pro Lys Ala
 290 295 300
 Leu Lys His Ala Gly Ile Glu Asp Ile Asn Ser Val Asp Tyr Phe Glu
 305 310 315 320
 Phe Asn Glu Ala Phe Ser Val Val Gly Leu Val Asn Thr Lys Ile Leu
 325 330 335
 Lys Leu Asp Pro Ser Lys Val Asn Val Tyr Gly Gly Ala Val Ala Leu
 340 345 350
 Gly His Pro Leu Gly Cys Ser Gly Ala Arg Val Val Val Thr Leu Leu
 355 360 365
 Ser Ile Leu Gln Gln Glu Gly Gly Lys Ile Gly Val Ala Ala Ile Cys
 370 375 380
 Asn Gly Gly Gly Gly Ala Ser Ser Ile Val Ile Glu Lys Ile
 385 390 395

<210> SEQ ID NO 134

<211> LENGTH: 864

<212> TYPE: DNA

<213> ORGANISM: *Bacillus subtilis*

<400> SEQUENCE: 134

```

atggaaatca aacaaatcat ggtagctggc gcaggtcaga tggggagcgg aattgctcaa    60
acagccgcgc acgcgggctt ttatgtgcgg atgtatgatg tgaatccaga ggccgcggag    120
gcaggattga aacggctgaa gaaacagctg gcccgatgat ctgagaaaagg aaaaaggacc    180
gagacggaag tgaagagcgt aatcaaccgc atttcgattt ctcaaacact tgaggaggca    240
gagcatgcgg acatttgtgat tgaggctatc gcagaaaaca tggcggcaaa aactgagatg    300
tttaaaacac ttgatcgcat ttgccgcct catatcgattt tggccagcaa tacatcttcc    360
ttgcctatta cagaaatcgc tgcgtgaaca aaccggcctc aacgggttat tgcatgcat    420
tttatgaatc cgtccctgt aatgaagctg gtagaagtga ttcgaggctt ggctacatca    480
gaagaaacgg ccttagatgt tatggcatta gcggaaaaga tggggaaaac agcggtagaa    540
gtcaatgatt ttcttgggtt tgtttccaac cgtgtgcttc ttccaatgat taatgaagcc    600
atctattgcg tgtatgaggg agtggcgaag ccggaggcaa tagatgaagt gatgaagctg    660
ggcatgaatc atccgatggg tccgcttgca ttageggatt ttatcgact ggatacgtgt    720
ttatcaatta tggaagtctt tcaactcaggc cttggcgatt ccaaataccg tccttgcccg    780
ctgctccgca agtatgtcaa agcaggtcgg cttggcaaaa agagcggacg cggtttttat    840
gactatgagg agaagacttc ctga                                     864
  
```

<210> SEQ ID NO 135

<211> LENGTH: 287

<212> TYPE: PRT

-continued

<213> ORGANISM: *Bacillus subtilis*

<400> SEQUENCE: 135

Met Glu Ile Lys Gln Ile Met Val Ala Gly Ala Gly Gln Met Gly Ser
 1 5 10 15
 Gly Ile Ala Gln Thr Ala Ala Asp Ala Gly Phe Tyr Val Arg Met Tyr
 20 25 30
 Asp Val Asn Pro Glu Ala Ala Glu Ala Gly Leu Lys Arg Leu Lys Lys
 35 40 45
 Gln Leu Ala Arg Asp Ala Glu Lys Gly Lys Arg Thr Glu Thr Glu Val
 50 55 60
 Lys Ser Val Ile Asn Arg Ile Ser Ile Ser Gln Thr Leu Glu Glu Ala
 65 70 75 80
 Glu His Ala Asp Ile Val Ile Glu Ala Ile Ala Glu Asn Met Ala Ala
 85 90 95
 Lys Thr Glu Met Phe Lys Thr Leu Asp Arg Ile Cys Pro Pro His Thr
 100 105 110
 Ile Leu Ala Ser Asn Thr Ser Ser Leu Pro Ile Thr Glu Ile Ala Ala
 115 120 125
 Val Thr Asn Arg Pro Gln Arg Val Ile Gly Met His Phe Met Asn Pro
 130 135 140
 Val Pro Val Met Lys Leu Val Glu Val Ile Arg Gly Leu Ala Thr Ser
 145 150 155 160
 Glu Glu Thr Ala Leu Asp Val Met Ala Leu Ala Glu Lys Met Gly Lys
 165 170 175
 Thr Ala Val Glu Val Asn Asp Phe Pro Gly Phe Val Ser Asn Arg Val
 180 185 190
 Leu Leu Pro Met Ile Asn Glu Ala Ile Tyr Cys Val Tyr Glu Gly Val
 195 200 205
 Ala Lys Pro Glu Ala Ile Asp Glu Val Met Lys Leu Gly Met Asn His
 210 215 220
 Pro Met Gly Pro Leu Ala Leu Ala Asp Phe Ile Gly Leu Asp Thr Cys
 225 230 235 240
 Leu Ser Ile Met Glu Val Leu His Ser Gly Leu Gly Asp Ser Lys Tyr
 245 250 255
 Arg Pro Cys Pro Leu Leu Arg Lys Tyr Val Lys Ala Gly Trp Leu Gly
 260 265 270
 Lys Lys Ser Gly Arg Gly Phe Tyr Asp Tyr Glu Glu Lys Thr Ser
 275 280 285

<210> SEQ ID NO 136

<211> LENGTH: 855

<212> TYPE: DNA

<213> ORGANISM: *Ralstonia eutropha*

<400> SEQUENCE: 136

atggcaatca ggacagtggg catcgtgggt gccggcacca tgggcaacgg catcgcgag 60
 gcttggtcgg tggtgggcct ggacgtggtg atggtggata tcagcgacgc agcgggtcag 120
 aagggcatcg ccaccgtcgc cggcagcctg gaccgcctga tcaagaagga caagatcagc 180
 gaagccgaca agatgactgc gctcgcgcgc atccacggca gcaccgcgta tgacgacctg 240
 aagaaggccg atatcgtgat cgaggccgcc accgagaact ttgacctgaa ggtcaagatc 300
 ctcaagcaga tcgacagcat cgtcggcgag aacgtcatca ttgcttcgaa cagtcgctcg 360
 atctcgatca ccaagctggc cgccgtgacg agtcgccccg agcgcttcac cggaatgcac 420

-continued

```

ttcttcaacc cgggtgccgt gatggcgctg gtggaactga tccgcggcct gcagaccagc 480
gacgcggctc acgccgatgt cgaggcgctg gccaaaggaac tgggcaagta cccgatcacc 540
gtcaagaaca gcccgggctt cgtcgtcaac cgcatacctgt gcccgatgat caacgaagcc 600
ttctgcgtgc tcggtgaagg cctggcctcg ccggaagaga tcgacgaagg catgaagctc 660
ggctgcaacc atccgatcgg cccctggca ctggccgaca tgateggcct ggacaccatg 720
ctggcagtga tggaagtgt gtacacagaa ttgcccgcgc cgaagtatcg tccggccatg 780
ctgatgcgcg agatggtggc tgcgggggtat ctgggccgca gaactggccg tggcgtgtac 840
gtctacagca agtaa 855

```

<210> SEQ ID NO 137

<211> LENGTH: 284

<212> TYPE: PRT

<213> ORGANISM: *Ralstonia eutropha*

<400> SEQUENCE: 137

```

Met Ala Ile Arg Thr Val Gly Ile Val Gly Ala Gly Thr Met Gly Asn
1           5           10          15
Gly Ile Ala Gln Ala Cys Ala Val Val Gly Leu Asp Val Val Met Val
20          25          30
Asp Ile Ser Asp Ala Ala Val Gln Lys Gly Ile Ala Thr Val Ala Gly
35          40          45
Ser Leu Asp Arg Leu Ile Lys Lys Asp Lys Ile Ser Glu Ala Asp Lys
50          55          60
Met Thr Ala Leu Ala Arg Ile His Gly Ser Thr Ala Tyr Asp Asp Leu
65          70          75          80
Lys Lys Ala Asp Ile Val Ile Glu Ala Ala Thr Glu Asn Phe Asp Leu
85          90          95
Lys Val Lys Ile Leu Lys Gln Ile Asp Ser Ile Val Gly Glu Asn Val
100         105        110
Ile Ile Ala Ser Asn Thr Ser Ser Ile Ser Ile Thr Lys Leu Ala Ala
115        120        125
Val Thr Ser Arg Pro Glu Arg Phe Ile Gly Met His Phe Phe Asn Pro
130        135        140
Val Pro Val Met Ala Leu Val Glu Leu Ile Arg Gly Leu Gln Thr Ser
145        150        155        160
Asp Ala Ala His Ala Asp Val Glu Ala Leu Ala Lys Glu Leu Gly Lys
165        170        175
Tyr Pro Ile Thr Val Lys Asn Ser Pro Gly Phe Val Val Asn Arg Ile
180        185        190
Leu Cys Pro Met Ile Asn Glu Ala Phe Cys Val Leu Gly Glu Gly Leu
195        200        205
Ala Ser Pro Glu Glu Ile Asp Glu Gly Met Lys Leu Gly Cys Asn His
210        215        220
Pro Ile Gly Pro Leu Ala Leu Ala Asp Met Ile Gly Leu Asp Thr Met
225        230        235        240
Leu Ala Val Met Glu Val Leu Tyr Thr Glu Phe Ala Asp Pro Lys Tyr
245        250        255
Arg Pro Ala Met Leu Met Arg Glu Met Val Ala Ala Gly Tyr Leu Gly
260        265        270
Arg Lys Thr Gly Arg Gly Val Tyr Val Tyr Ser Lys
275        280

```

-continued

<210> SEQ ID NO 138
 <211> LENGTH: 741
 <212> TYPE: DNA
 <213> ORGANISM: Alcaligenes eutrophis

<400> SEQUENCE: 138

```

atgactcagc gcattgcgta tgtgaccggc ggcattgggtg gtatcggaac cgccatttgc    60
cagcggttgg ccaaggatgg ctttcgtgtg gtggccggtt gcggccccaa ctgcgcgcgc    120
cgcgaaaagt ggctggagca gcagaaggcc ctgggcttcg atttcattgc ctcggaaggc    180
aatgtggctg actgggactc gaccaagacc gcattcgaca aggtcaagtc cgaggtcggc    240
gaggttgatg tgctgatcaa caacgccggt atcaccgcgc acgtggtgtt ccgcaagatg    300
acccgcgcgc actgggatgc ggtgatcgac accaacctga cctcgctgtt caacgtcacc    360
aagcaggtga tcgacggcat ggccgaccgt ggctggggcc gcatcgtaaa catctcgtcg    420
gtgaacgggc agaagggcca gttcgccag accaactact ccaccgcaa ggccggcctg    480
catggcttca ccatggcact ggccgaggaa gtggcgacca agggcgtgac cgtcaacacg    540
gtctctccgg gctatatcgc caccgacatg gtcaaggcga tccgccagga cgtgctcgac    600
aagatcgtcg cgacgatccc ggtcaagcgc ctgggcctgc cggaagagat cgctcgatc    660
tgccgctggt tgctgctcga ggagtcggt ttctcgaccg gcgccgactt ctgctcaac    720
ggcggcctgc atatgggctg a                                         741

```

<210> SEQ ID NO 139
 <211> LENGTH: 246
 <212> TYPE: PRT
 <213> ORGANISM: Alcaligenes eutrophus

<400> SEQUENCE: 139

```

Met Thr Gln Arg Ile Ala Tyr Val Thr Gly Gly Met Gly Gly Ile Gly
1           5           10          15
Thr Ala Ile Cys Gln Arg Leu Ala Lys Asp Gly Phe Arg Val Val Ala
20          25          30
Gly Cys Gly Pro Asn Ser Pro Arg Arg Glu Lys Trp Leu Glu Gln Gln
35          40          45
Lys Ala Leu Gly Phe Asp Phe Ile Ala Ser Glu Gly Asn Val Ala Asp
50          55          60
Trp Asp Ser Thr Lys Thr Ala Phe Asp Lys Val Lys Ser Glu Val Gly
65          70          75          80
Glu Val Asp Val Leu Ile Asn Asn Ala Gly Ile Thr Arg Asp Val Val
85          90          95
Phe Arg Lys Met Thr Arg Ala Asp Trp Asp Ala Val Ile Asp Thr Asn
100         105         110
Leu Thr Ser Leu Phe Asn Val Thr Lys Gln Val Ile Asp Gly Met Ala
115         120         125
Asp Arg Gly Trp Gly Arg Ile Val Asn Ile Ser Ser Val Asn Gly Gln
130         135         140
Lys Gly Gln Phe Gly Gln Thr Asn Tyr Ser Thr Ala Lys Ala Gly Leu
145         150         155         160
His Gly Phe Thr Met Ala Leu Ala Gln Glu Val Ala Thr Lys Gly Val
165         170         175
Thr Val Asn Thr Val Ser Pro Gly Tyr Ile Ala Thr Asp Met Val Lys
180         185         190

```

[illegible]

```
<210> SEQ ID NO 140
<211> LENGTH: 768
<212> TYPE: DNA
<213> ORGANISM: Escherichia coli
```

<400> SEQUENCE: 140

atgagcgaac	tgatcgctag	cgcgcagcaa	cgagtattgt	tgctgacctt	taaccgtccc	60
gccgcacgta	atgcgctaaa	taatgcacctg	ctgatgcaac	tggtaaatga	actggaagct	120
gcggctaccg	ataccagcat	ttcggctctgt	gtgattaccg	gtaatgcacg	ctttttttgcc	180
gctggggcgcg	atctcaacga	aatggcagaa	aaagatctcg	cggccacctt	aaacgataca	240
cgtccgcagc	tatgggcgcg	attgcaggcc	tttaataaac	cacttatcgc	agccgtcaat	300
ggttacgcgc	ttggggcggg	ttgcgaactg	gcattgttgt	gcgatgtggt	ggttgcggga	360
gagaacgcgc	gttttggggt	gccggaaaac	actctcggca	tcatgcctgg	cgcaggcgga	420
acgcaacgtt	taatccgtag	tgctcggtaaa	tcgttagcca	gcaaaatggt	gctgagcgga	480
gaaagtatca	ccgctcagca	agcacagcag	gccgggctgg	ttagcgacgt	cttcccacgc	540
gatttaaccc	tcgaatacgc	cttacagctg	gcatacga	tgggcacgtca	ctcgccgctg	600
gccttacaag	cggcaaaagca	agcgcctgcg	cagtcgcagg	aagtggcttt	gcaagccgga	660
cttgcccagg	agcgacagtt	attcaccttg	ctggcggcaa	cagaagatcg	tcatgaaggc	720
atctccgctt	tcttacaaaa	acgcacgcgc	gacttttaa	gacgctaa		768

```
<210> SEQ ID NO 141
<211> LENGTH: 255
<212> TYPE: PRT
<213> ORGANISM: Escherichia coli
```

<400> SEQUENCE: 141

Met	Ser	Glu	Leu	Ile	Val	Ser	Arg	Gln	Gln	Arg	Val	Leu	Leu	Leu	Thr
1				5					10					15	
Leu	Asn	Arg	Pro	Ala	Ala	Arg	Asn	Ala	Leu	Asn	Asn	Ala	Leu	Leu	Met
			20					25					30		
Gln	Leu	Val	Asn	Glu	Leu	Glu	Ala	Ala	Ala	Thr	Asp	Thr	Ser	Ile	Ser
		35					40					45			
Val	Cys	Val	Ile	Thr	Gly	Asn	Ala	Arg	Phe	Phe	Ala	Ala	Gly	Ala	Asp
	50					55					60				
Leu	Asn	Glu	Met	Ala	Glu	Lys	Asp	Leu	Ala	Ala	Thr	Leu	Asn	Asp	Thr
65					70					75					80
Arg	Pro	Gln	Leu	Trp	Ala	Arg	Leu	Gln	Ala	Phe	Asn	Lys	Pro	Leu	Ile
				85					90					95	
Ala	Ala	Val	Asn	Gly	Tyr	Ala	Leu	Gly	Ala	Gly	Cys	Glu	Leu	Ala	Leu
			100					105					110		
Leu	Cys	Asp	Val	Val	Val	Ala	Gly	Glu	Asn	Ala	Arg	Phe	Gly	Leu	Pro
		115					120					125			

-continued

Glu Ile Thr Leu Gly Ile Met Pro Gly Ala Gly Gly Thr Gln Arg Leu
 130 135 140

Ile Arg Ser Val Gly Lys Ser Leu Ala Ser Lys Met Val Leu Ser Gly
 145 150 155 160

Glu Ser Ile Thr Ala Gln Gln Ala Gln Gln Ala Gly Leu Val Ser Asp
 165 170 175

Val Phe Pro Ser Asp Leu Thr Leu Glu Tyr Ala Leu Gln Leu Ala Ser
 180 185 190

Lys Met Ala Arg His Ser Pro Leu Ala Leu Gln Ala Ala Lys Gln Ala
 195 200 205

Leu Arg Gln Ser Gln Glu Val Ala Leu Gln Ala Gly Leu Ala Gln Glu
 210 215 220

Arg Gln Leu Phe Thr Leu Leu Ala Ala Thr Glu Asp Arg His Glu Gly
 225 230 235 240

Ile Ser Ala Phe Leu Gln Lys Arg Thr Pro Asp Phe Lys Gly Arg
 245 250 255

<210> SEQ ID NO 142
 <211> LENGTH: 783
 <212> TYPE: DNA
 <213> ORGANISM: Bacillus subtilis

<400> SEQUENCE: 142

atgggagatt ctattctttt tactgttaaa aatgaacata tggcgttgat caccttaaac	60
aggcctcagg cagcaaatgc tctttcagcg gaaatgctta gaaacctgca aatgattatc	120
caggaaattg aatttaactc aaacatccgt tgcgtcatcc tcacaggcac cggtgaaaaa	180
gcgtttttgtg caggggcaga cctgaaggaa cggataaaac tgaaagaaga tcaggttctg	240
gaaagtgtat ctctcattca aagaacggcg gctttacttg atgccttgcc gcagccggtc	300
atagctgcga taaatggaag cgcattagcg gccggactag aattggcatt ggcatgcgac	360
cttcgaatcg caactgaagc agctgtgctg ggacttccgg aaacagggtt agctattatc	420
ccggcgctg gagggaccca aaggctgccc cggctgattg gcagaggaaa agcaaaagaa	480
ttcatttata caggcagacg cgtgaccgca cacgaagcaa aagaaatcgg ccttgtagag	540
catgtcacgg ctcccttgta ccttatgcc aagcagagg aactggccgc agccatttct	600
gccaacggac cgatcgctgt ccgtcaggct aaatttgcaa tcaataaagg attggagaca	660
gatcttgcta caggccttgc gattgaacaa aaagcgtatg aacaaacat cccgacaaaa	720
gacaggagag aagggcttca ggcctttcaa gaaaaagac gggccgtata caagggaata	780
taa	783

<210> SEQ ID NO 143
 <211> LENGTH: 260
 <212> TYPE: PRT
 <213> ORGANISM: Bacillus subtilis

<400> SEQUENCE: 143

Met Gly Asp Ser Ile Leu Phe Thr Val Lys Asn Glu His Met Ala Leu
 1 5 10 15

Ile Thr Leu Asn Arg Pro Gln Ala Ala Asn Ala Leu Ser Ala Glu Met
 20 25 30

Leu Arg Asn Leu Gln Met Ile Ile Gln Glu Ile Glu Phe Asn Ser Asn
 35 40 45

Ile Arg Cys Val Ile Leu Thr Gly Thr Gly Glu Lys Ala Phe Cys Ala
 50 55 60

-continued

Gly Ala Asp Leu Lys Glu Arg Ile Lys Leu Lys Glu Asp Gln Val Leu
 65 70 75 80
 Glu Ser Val Ser Leu Ile Gln Arg Thr Ala Ala Leu Leu Asp Ala Leu
 85 90 95
 Pro Gln Pro Val Ile Ala Ala Ile Asn Gly Ser Ala Leu Gly Gly Gly
 100 105 110
 Leu Glu Leu Ala Leu Ala Cys Asp Leu Arg Ile Ala Thr Glu Ala Ala
 115 120 125
 Val Leu Gly Leu Pro Glu Thr Gly Leu Ala Ile Ile Pro Gly Ala Gly
 130 135 140
 Gly Thr Gln Arg Leu Pro Arg Leu Ile Gly Arg Gly Lys Ala Lys Glu
 145 150 155 160
 Phe Ile Tyr Thr Gly Arg Arg Val Thr Ala His Glu Ala Lys Glu Ile
 165 170 175
 Gly Leu Val Glu His Val Thr Ala Pro Cys Asp Leu Met Pro Lys Ala
 180 185 190
 Glu Glu Leu Ala Ala Ala Ile Ser Ala Asn Gly Pro Ile Ala Val Arg
 195 200 205
 Gln Ala Lys Phe Ala Ile Asn Lys Gly Leu Glu Thr Asp Leu Ala Thr
 210 215 220
 Gly Leu Ala Ile Glu Gln Lys Ala Tyr Glu Gln Thr Ile Pro Thr Lys
 225 230 235 240
 Asp Arg Arg Glu Gly Leu Gln Ala Phe Gln Glu Lys Arg Arg Ala Val
 245 250 255
 Tyr Lys Gly Ile
 260

<210> SEQ ID NO 144
 <211> LENGTH: 405
 <212> TYPE: DNA
 <213> ORGANISM: *Aeromonas caviae*

<400> SEQUENCE: 144

```

atgagcgcac aatccctgga agtaggccag aaggcccgtc tcagcaagcg gttcggggcg      60
gcggaggttag ccgccttcgc cgcgtctctg gaggacttca accccctgca cctggaccgc      120
gccttcgcgc ccaccacggc gttcgagcgg cccatagtcc acggcatgct gctcgccagc      180
ctcttctccg ggctgctggg ccagcagttg ccgggcaagg ggagcatcta tctgggtcaa      240
agcctcagct tcaagctgcc ggtctttgtc ggggacgagg tgacggccga ggtggaggtg      300
accgcccttc gcgaggacaa gcccatcgcc accctgacca cccgcatctt cacccaaggc      360
ggcgccctcg ccgtgacggg ggaagccgtg gtcaagctgc cttaa                          405
  
```

<210> SEQ ID NO 145
 <211> LENGTH: 134
 <212> TYPE: PRT
 <213> ORGANISM: *Aeromonas caviae*

<400> SEQUENCE: 145

Met Ser Ala Gln Ser Leu Glu Val Gly Gln Lys Ala Arg Leu Ser Lys
 1 5 10 15
 Arg Phe Gly Ala Ala Glu Val Ala Ala Phe Ala Ala Leu Ser Glu Asp
 20 25 30
 Phe Asn Pro Leu His Leu Asp Pro Ala Phe Ala Ala Thr Thr Ala Phe
 35 40 45

-continued

Glu Arg Pro Ile Val His Gly Met Leu Leu Ala Ser Leu Phe Ser Gly
50 55 60

Leu Leu Gly Gln Gln Leu Pro Gly Lys Gly Ser Ile Tyr Leu Gly Gln
65 70 75 80

Ser Leu Ser Phe Lys Leu Pro Val Phe Val Gly Asp Glu Val Thr Ala
85 90 95

Glu Val Glu Val Thr Ala Leu Arg Glu Asp Lys Pro Ile Ala Thr Leu
100 105 110

Thr Thr Arg Ile Phe Thr Gln Gly Gly Ala Leu Ala Val Thr Gly Glu
115 120 125

Ala Val Val Lys Leu Pro
130

<210> SEQ ID NO 146
<211> LENGTH: 1912
<212> TYPE: DNA
<213> ORGANISM: Euglena gracilis

<400> SEQUENCE: 146

ttttcgcccg tgcaccacga tgcctgccc cgcctcgccg tctgctgccg tgggtgtctgc	60
cggcgccctc tgctgtgctg tggcaacggg attgttgccg actggatcca accccaccgc	120
cctgtccact gcttccactc gctctccgac ctactgggc cgtgggggtg acaggggctt	180
gatgaggcca accactgcag cggtctgac gacaatgaga gaggtgcccc agatggctga	240
gggattttca ggcaagcca cgtctgcatg ggcgcgcgcg gggcgcagc gggcggcgc	300
gctcgtggcc gcggcctcct ccgactggc gctgtgggtg tgggcgcgcc ggcgcagcgt	360
gcggcgccgc ctggcagcgc tggcggagct gccaccgcg gtcaccacc tggccccccc	420
gatggcgatg ttcaccacca cagcgaagg catccagccc aagattcgtg gcttcactg	480
cacgaccacc caccgatcg gctgtgagaa gcgggtccag gaggagatcg cgtacgccg	540
tgccccccg cccaccagcc ctggcccgaa gaggtgtctg gtcacggct gcagtaccg	600
ctacgggctc tccaccgcga tcaccgctgc ctccggctac caggccgcca cgtggggct	660
gttcttgccg gggccccga cgaagggcg ccccgccgcg gcgggctggg acaacaccgt	720
ggcgcttcag aaggccgccc tggaggccgg gctgtacgcc cggagcctta atggcgacgc	780
cttcgactcc acaacgaagg cgcggacggg cgaggcgatc aagcgggacc tcggcacggg	840
ggacctcgtg gtgtacagca tcgccgccc gaagcggacg gacctgcca ccggcgtcct	900
ccacaaggcc tgctgaagc ccacggcgc cacgtacacc aaccgcactg tgaacaccga	960
caaggcggag gtgaccgacg tcagcattga gccggcctcc cccgaagaga tcgcggacac	1020
ggtgaagggt atgggcgggg aggactggga gctctggatc caggcgctgt cggaggccgg	1080
cgtgctggcg gagggggcca agcgggtggc gtactcctac atcgccccg agatgacgtg	1140
gcctgtctac tggctccgga ccacggggga ggccaagaag gacgtggaga aggtgccaa	1200
gcgcatcacg cagcagtagc gctgcccggc gtaccgggtg gtggccaagg ccttggtcac	1260
ccaggccagc tccgccatcc cgggtgtgcc gctctacatc tgctgtgtgt accgcgttat	1320
gaaggagaag ggcacccacg agggctgcat cgagcagatg gtgcggctgc tcaccacgaa	1380
gctgtacccc gagaacgggg ccccatcgt cgatgaggcc ggacgtgtgc ggggtggatga	1440
ctgggagatg gcggaggatg tgcagcaggc tgtaaggac ctctggagcc aggtgagcac	1500
tgccaacctc aaggacatct ccgacttcgc tgggtatcaa actgagttcc tcgggtgtt	1560
cgggttcggc attgacggcg tggactacga ccagcccggtg gacgtggagg cggacctccc	1620

-continued

```

cagtgtgtgcc cagcagtagg tgctggacgc cgctctcttc cggggggtct gccaaaatgg 1680
tcgtctcccc aacccaaccc cctgcccacc atcggggtcc cgcggtgaa tgcggccccc 1740
acccaaaggc aaaggtcaag gccggggccc caccgccaaa gggtaacaca tatgtatccg 1800
tcgggggctg atccgctgc gacacgggcc ataattgtgc cccacgggat gtccatgcgc 1860
ctaagacaac tgccccggcc gacagtcgct accgccttga gttcccagg ca 1912

```

<210> SEQ ID NO 147

<211> LENGTH: 539

<212> TYPE: PRT

<213> ORGANISM: *Euglena gracilis*

<400> SEQUENCE: 147

```

Met Ser Cys Pro Ala Ser Pro Ser Ala Ala Val Val Ser Ala Gly Ala
1          5          10          15
Leu Cys Leu Cys Val Ala Thr Val Leu Leu Ala Thr Gly Ser Asn Pro
20          25          30
Thr Ala Leu Ser Thr Ala Ser Thr Arg Ser Pro Thr Ser Leu Val Arg
35          40          45
Gly Val Asp Arg Gly Leu Met Arg Pro Thr Thr Ala Ala Ala Leu Thr
50          55          60
Thr Met Arg Glu Val Pro Gln Met Ala Glu Gly Phe Ser Gly Glu Ala
65          70          75          80
Thr Ser Ala Trp Ala Ala Ala Gly Pro Gln Trp Ala Ala Pro Leu Val
85          90          95
Ala Ala Ala Ser Ser Ala Leu Ala Leu Trp Trp Trp Ala Ala Arg Arg
100         105         110
Ser Val Arg Arg Pro Leu Ala Ala Leu Ala Glu Leu Pro Thr Ala Val
115         120         125
Thr His Leu Ala Pro Pro Met Ala Met Phe Thr Thr Thr Ala Lys Val
130         135         140
Ile Gln Pro Lys Ile Arg Gly Phe Ile Cys Thr Thr Thr His Pro Ile
145         150         155         160
Gly Cys Glu Lys Arg Val Gln Glu Glu Ile Ala Tyr Ala Arg Ala His
165         170         175
Pro Pro Thr Ser Pro Gly Pro Lys Arg Val Leu Val Ile Gly Cys Ser
180         185         190
Thr Gly Tyr Gly Leu Ser Thr Arg Ile Thr Ala Ala Phe Gly Tyr Gln
195         200         205
Ala Ala Thr Leu Gly Val Phe Leu Ala Gly Pro Pro Thr Lys Gly Arg
210         215         220
Pro Ala Ala Ala Gly Trp Tyr Asn Thr Val Ala Phe Glu Lys Ala Ala
225         230         235         240
Leu Glu Ala Gly Leu Tyr Ala Arg Ser Leu Asn Gly Asp Ala Phe Asp
245         250         255
Ser Thr Thr Lys Ala Arg Thr Val Glu Ala Ile Lys Arg Asp Leu Gly
260         265         270
Thr Val Asp Leu Val Val Tyr Ser Ile Ala Ala Pro Lys Arg Thr Asp
275         280         285
Pro Ala Thr Gly Val Leu His Lys Ala Cys Leu Lys Pro Ile Gly Ala
290         295         300
Thr Tyr Thr Asn Arg Thr Val Asn Thr Asp Lys Ala Glu Val Thr Asp
305         310         315         320

```

Val	Ser	Ile	Glu	Pro	Ala	Ser	Pro	Glu	Glu	Ile	Ala	Asp	Thr	Val	Lys			
								325							330			335
Val	Met	Gly	Gly	Glu	Asp	Trp	Glu	Leu	Trp	Ile	Gln	Ala	Leu	Ser	Glu			
								340							345			350
Ala	Gly	Val	Leu	Ala	Glu	Gly	Ala	Lys	Thr	Val	Ala	Tyr	Ser	Tyr	Ile			
								355							360			365
Gly	Pro	Glu	Met	Thr	Trp	Pro	Val	Tyr	Trp	Ser	Gly	Thr	Ile	Gly	Glu			
								370							375			380
Ala	Lys	Lys	Asp	Val	Glu	Lys	Ala	Ala	Lys	Arg	Ile	Thr	Gln	Gln	Tyr			
								385							390			400
Gly	Cys	Pro	Ala	Tyr	Pro	Val	Val	Ala	Lys	Ala	Leu	Val	Thr	Gln	Ala			
								405							410			415
Ser	Ser	Ala	Ile	Pro	Val	Val	Pro	Leu	Tyr	Ile	Cys	Leu	Leu	Tyr	Arg			
								420							425			430
Val	Met	Lys	Glu	Lys	Gly	Thr	His	Glu	Gly	Cys	Ile	Glu	Gln	Met	Val			
								435							440			445
Arg	Leu	Leu	Thr	Thr	Lys	Leu	Tyr	Pro	Glu	Asn	Gly	Ala	Pro	Ile	Val			
								450							455			460
Asp	Glu	Ala	Gly	Arg	Val	Arg	Val	Asp	Asp	Trp	Glu	Met	Ala	Glu	Asp			
								465							470			475
Val	Gln	Gln	Ala	Val	Lys	Asp	Leu	Trp	Ser	Gln	Val	Ser	Thr	Ala	Asn			
								485							490			495
Leu	Lys	Asp	Ile	Ser	Asp	Phe	Ala	Gly	Tyr	Gln	Thr	Glu	Phe	Leu	Arg			
								500							505			510
Leu	Phe	Gly	Phe	Gly	Ile	Asp	Gly	Val	Asp	Tyr	Asp	Gln	Pro	Val	Asp			
								515							520			525
Val	Glu	Ala	Asp	Leu	Pro	Ser	Ala	Ala	Gln	Gln								
								530							535			

```
<210> SEQ ID NO 148
<211> LENGTH: 1344
<212> TYPE: DNA
<213> ORGANISM: Streptomyces collinus

<400> SEQUENCE: 148
```

gtgaccgtga	aggacatctc	ggacgcgac	cagtcgaagg	acgccacgtc	cgccgacttc	60
gccgccctgc	agctcccccga	gtcgtaccgt	gcgatcaccg	tgcacaagga	cgagacggag	120
atgttcgcgg	gtctggagac	ccgcgacaag	gaccgcgcga	agtcgatcca	cctcgacgag	180
gtgccctgtc	ccgaactggg	cccgggcgaa	gcccctggtg	ccgtcatggc	ctcctcggtc	240
aactacaact	cggtgtggac	ctcgatcttc	gagccggtgt	cgacgttcgc	cttcttgag	300
cgctacggca	agctgtcgcc	gctgaccaag	cgccacgacc	tgccgtacca	catcatcggc	360
tccgacctcg	cgggcgctgt	cctgcgcacc	ggccccggcg	tcaacgcctg	gcagcccggt	420
gacgaggctc	tcgcgcactg	cctgagcgtc	gagctggagt	cgcccgacgg	ccacgacgac	480
accatgctcg	accccgagca	gcgcacatcg	ggcttcgaga	ccaacttcgg	cggcctcgcg	540
gagatcgcg	tggtcaagac	gaaccagctg	atgccgaagc	cgaagcacct	cacctgggag	600
gaggcccgcg	ccccgggcct	ggtgaactcc	accgcctacc	gccagctggt	ctcccgaac	660
ggcgccgcga	tgaagcaggg	gcacaacgtc	ctgatctggg	gcgcgagcgg	cgggctcggc	720
tcgtacgcca	cgcagttcgc	gctcgcgggc	ggtgccaaac	cgatctgtgt	cgtctcctcg	780
ccccagaagg	cggagatctg	ccgctcgatg	ggcgccgagg	cgatcatcga	ccgcaacgcc	840
gagggctaca	agttctggaa	ggacgagcac	accaggacc	ccaaggagtg	gaagcgcttc	900

-continued

```

ggcaagcgca tccgcgagct gaccggcggc gaggacatcg acatcgtctt cgagcaccct 960
ggccgcgaga ccttcgggcg ctcctgtctac gtcacccgca agggcggcac catcaccacc 1020
tgcgccctga cctcgggcta catgcacgag tacgacaacc ggtacctgtg gatgtccctg 1080
aagcggatca tcggtctgca cttcgccaac taccgcgagg cgtagcaggg caaccgctg 1140
atcgccaagg gcaagatcca cccgacgtg tcgaagacgt actccttgga ggagaccggc 1200
caggcggcgt acgacgtcca ccgcaacctg caccagggca aggtcggcgt cctgtgctc 1260
gcgcgggagg aaggcctcgg cgtgcgcgac gcggagatgc gcgccagca catcgacgcc 1320
atcaaccgct tccgcaacgt ctga 1344

```

<210> SEQ ID NO 149

<211> LENGTH: 447

<212> TYPE: PRT

<213> ORGANISM: Streptomyces collinus

<400> SEQUENCE: 149

```

Met Thr Val Lys Asp Ile Leu Asp Ala Ile Gln Ser Lys Asp Ala Thr
1           5           10          15
Ser Ala Asp Phe Ala Ala Leu Gln Leu Pro Glu Ser Tyr Arg Ala Ile
20          25          30
Thr Val His Lys Asp Glu Thr Glu Met Phe Ala Gly Leu Glu Thr Arg
35          40          45
Asp Lys Asp Pro Arg Lys Ser Ile His Leu Asp Glu Val Pro Val Pro
50          55          60
Glu Leu Gly Pro Gly Glu Ala Leu Val Ala Val Met Ala Ser Ser Val
65          70          75          80
Asn Tyr Asn Ser Val Trp Thr Ser Ile Phe Glu Pro Val Ser Thr Phe
85          90          95
Ala Phe Leu Glu Arg Tyr Gly Lys Leu Ser Pro Leu Thr Lys Arg His
100         105         110
Asp Leu Pro Tyr His Ile Ile Gly Ser Asp Leu Ala Gly Val Val Leu
115         120         125
Arg Thr Gly Pro Gly Val Asn Ala Trp Gln Pro Gly Asp Glu Val Val
130         135         140
Ala His Cys Leu Ser Val Glu Leu Glu Ser Pro Asp Gly His Asp Asp
145         150         155         160
Thr Met Leu Asp Pro Glu Gln Arg Ile Trp Gly Phe Glu Thr Asn Phe
165         170         175
Gly Gly Leu Ala Glu Ile Ala Leu Val Lys Thr Asn Gln Leu Met Pro
180         185         190
Lys Pro Lys His Leu Thr Trp Glu Glu Ala Ala Ala Pro Gly Leu Val
195         200         205
Asn Ser Thr Ala Tyr Arg Gln Leu Val Ser Arg Asn Gly Ala Ala Met
210         215         220
Lys Gln Gly Asp Asn Val Leu Ile Trp Gly Ala Ser Gly Gly Leu Gly
225         230         235         240
Ser Tyr Ala Thr Gln Phe Ala Leu Ala Gly Gly Ala Asn Pro Ile Cys
245         250         255
Val Val Ser Ser Pro Gln Lys Ala Glu Ile Cys Arg Ser Met Gly Ala
260         265         270
Glu Ala Ile Ile Asp Arg Asn Ala Glu Gly Tyr Lys Phe Trp Lys Asp
275         280         285

```

-continued

Glu His Thr Gln Asp Pro Lys Glu Trp Lys Arg Phe Gly Lys Arg Ile
 290 295 300
 Arg Glu Leu Thr Gly Gly Glu Asp Ile Asp Ile Val Phe Glu His Pro
 305 310 315 320
 Gly Arg Glu Thr Phe Gly Ala Ser Val Tyr Val Thr Arg Lys Gly Gly
 325 330 335
 Thr Ile Thr Thr Cys Ala Ser Thr Ser Gly Tyr Met His Glu Tyr Asp
 340 345 350
 Asn Arg Tyr Leu Trp Met Ser Leu Lys Arg Ile Ile Gly Ser His Phe
 355 360 365
 Ala Asn Tyr Arg Glu Ala Tyr Glu Ala Asn Arg Leu Ile Ala Lys Gly
 370 375 380
 Lys Ile His Pro Thr Leu Ser Lys Thr Tyr Ser Leu Glu Glu Thr Gly
 385 390 395 400
 Gln Ala Ala Tyr Asp Val His Arg Asn Leu His Gln Gly Lys Val Gly
 405 410 415
 Val Leu Cys Leu Ala Pro Glu Glu Gly Leu Gly Val Arg Asp Ala Glu
 420 425 430
 Met Arg Ala Gln His Ile Asp Ala Ile Asn Arg Phe Arg Asn Val
 435 440 445

<210> SEQ ID NO 150

<211> LENGTH: 1344

<212> TYPE: DNA

<213> ORGANISM: Streptomyces coelicolor

<400> SEQUENCE: 150

```

gtgaccgtga aggacatcct ggacgcgatc cagtcgcccc actccacgcc ggccgacatc      60
gccgcactgc cgctccccga gtcgtaccgc gcgatcacgc tgcacaagga cgagaccgag      120
atgttcgcgg gcctcgagac ccgcgacaag gacccccgca agtcgatcca cctggacgac      180
gtgccggtgc ccgagctggg ccccgcgagc gccctgggtg ccgtcatggc ctccctcggtc      240
aactacaact cgggtgtggc ctcgatcttc gagccgctgt ccaccttcgg gttcctggag      300
cgctacggcc gggtcagcga cctcgccaag cggcacgacc tgccgtacca cgtcacggc      360
tccgacctcg ccggtgtcgt cctgcgcacc ggtccggcgc tcaacgcctg gcaggcgggc      420
gacgaggctg tcgcgcactg cctctccgtc gagctggagt cctccgacgg ccacaacgac      480
acgatgctcg accccgagca gcgcattctg ggcttcgaga ccaacttcgg cggcctcgcg      540
gagatcgcgc tggtcaagtc caaccagctg atgccgaagc cggaccacct gagctgggag      600
gaggccgccc ctcccggcct ggtcaactcc accgcgtacc gccagctcgt ctcccgaac      660
ggcgccggca tgaagcaggg cgacaacgtg ctcatctggg gcgcgagcgg cggactcggc      720
tcgtacgcca ccagttcgc cctcgccggc ggccccaacc cgatctgcgt cgtctcctcg      780
ccgcagaagg cggagatctg ccgcgcgatg ggccgcgagg cgatcatcga ccgcaacgcc      840
gagggctacc ggttctggaa ggacgagaa acccaggacc cgaaggagtg gaagcgcttc      900
ggcaagcgca tccgcgaact gaccggcggc gaggacatcg acatcgtctt cgagcaaccc      960
ggccgcgaga ccttcggcgc ctccgtcttc gtcacccgca agggcggcac catcaccacc     1020
tgcgccctga cctcgggcta catgcacgag tacgacaacc gctacctgtg gatgtccctg     1080
aagcgcatca tcggctcgca cttcgccaac taccgcgagg cctgggaggc caaccgctc     1140
atgcccaagg gcaggatcca cccacgctc tccaaggtgt actccctcga ggacaccggc     1200
caggccgcct acgacgtcca ccgcaacctc caccagggca aggtcggcgt gctgtgctcg     1260

```

-continued

gagcccgagg agggcctggg cgtgcgcgac cgggagaagc gcgcgcagca cctcgacgcc 1320

atcaaccgct tccggaacat ctga 1344

<210> SEQ ID NO 151

<211> LENGTH: 447

<212> TYPE: PRT

<213> ORGANISM: Streptomyces coelicolor

<400> SEQUENCE: 151

Met Thr Val Lys Asp Ile Leu Asp Ala Ile Gln Ser Pro Asp Ser Thr
1 5 10 15

Pro Ala Asp Ile Ala Ala Leu Pro Leu Pro Glu Ser Tyr Arg Ala Ile
20 25 30

Thr Val His Lys Asp Glu Thr Glu Met Phe Ala Gly Leu Glu Thr Arg
35 40 45

Asp Lys Asp Pro Arg Lys Ser Ile His Leu Asp Asp Val Pro Val Pro
50 55 60

Glu Leu Gly Pro Gly Glu Ala Leu Val Ala Val Met Ala Ser Ser Val
65 70 75 80

Asn Tyr Asn Ser Val Trp Thr Ser Ile Phe Glu Pro Leu Ser Thr Phe
85 90 95

Gly Phe Leu Glu Arg Tyr Gly Arg Val Ser Asp Leu Ala Lys Arg His
100 105 110

Asp Leu Pro Tyr His Val Ile Gly Ser Asp Leu Ala Gly Val Val Leu
115 120 125

Arg Thr Gly Pro Gly Val Asn Ala Trp Gln Ala Gly Asp Glu Val Val
130 135 140

Ala His Cys Leu Ser Val Glu Leu Glu Ser Ser Asp Gly His Asn Asp
145 150 155 160

Thr Met Leu Asp Pro Glu Gln Arg Ile Trp Gly Phe Glu Thr Asn Phe
165 170 175

Gly Gly Leu Ala Glu Ile Ala Leu Val Lys Ser Asn Gln Leu Met Pro
180 185 190

Lys Pro Asp His Leu Ser Trp Glu Glu Ala Ala Ala Pro Gly Leu Val
195 200 205

Asn Ser Thr Ala Tyr Arg Gln Leu Val Ser Arg Asn Gly Ala Gly Met
210 215 220

Lys Gln Gly Asp Asn Val Leu Ile Trp Gly Ala Ser Gly Gly Leu Gly
225 230 235 240

Ser Tyr Ala Thr Gln Phe Ala Leu Ala Gly Gly Ala Asn Pro Ile Cys
245 250 255

Val Val Ser Ser Pro Gln Lys Ala Glu Ile Cys Arg Ala Met Gly Ala
260 265 270

Glu Ala Ile Ile Asp Arg Asn Ala Glu Gly Tyr Arg Phe Trp Lys Asp
275 280 285

Glu Asn Thr Gln Asp Pro Lys Glu Trp Lys Arg Phe Gly Lys Arg Ile
290 295 300

Arg Glu Leu Thr Gly Gly Glu Asp Ile Asp Ile Val Phe Glu His Pro
305 310 315 320

Gly Arg Glu Thr Phe Gly Ala Ser Val Phe Val Thr Arg Lys Gly Gly
325 330 335

Thr Ile Thr Thr Cys Ala Ser Thr Ser Gly Tyr Met His Glu Tyr Asp
340 345 350

-continued

Asn	Arg	Tyr	Leu	Trp	Met	Ser	Leu	Lys	Arg	Ile	Ile	Gly	Ser	His	Phe
		355					360					365			
Ala	Asn	Tyr	Arg	Glu	Ala	Trp	Glu	Ala	Asn	Arg	Leu	Ile	Ala	Lys	Gly
	370					375					380				
Arg	Ile	His	Pro	Thr	Leu	Ser	Lys	Val	Tyr	Ser	Leu	Glu	Asp	Thr	Gly
	385				390					395					400
Gln	Ala	Ala	Tyr	Asp	Val	His	Arg	Asn	Leu	His	Gln	Gly	Lys	Val	Gly
			405						410					415	
Val	Leu	Cys	Leu	Ala	Pro	Glu	Glu	Gly	Leu	Gly	Val	Arg	Asp	Arg	Glu
			420					425					430		
Lys	Arg	Ala	Gln	His	Leu	Asp	Ala	Ile	Asn	Arg	Phe	Arg	Asn	Ile	
		435					440						445		

<210> SEQ ID NO 152

<211> LENGTH: 2589

<212> TYPE: DNA

<213> ORGANISM: Clostridium acetobutylicum

<400> SEQUENCE: 152

atgaaagtca caacagtaaa ggaattagat gaaaaactca aggtaattaa agaagctcaa	60
aaaaaattct cttgttactc gcaagaaatg gttgatgaaa tctttagaaa tgcagcaatg	120
gcagcaatcg acgcaaggat agagctagca aaagcagctg ttttgaaac cggtatgggc	180
ttagttgaag acaaggttat aaaaaatcat ttgcaggcg aatacatcta taacaaatat	240
aaggatgaaa aaacctgagg tataattgaa cgaaatgaac cctacggaat tacaaaaata	300
gcagaaccta taggagttgt agctgtctata atccctgtaa caaacccac atcaacaaca	360
atatttaaat ccttaatatc ccttaaaact agaaatggaa tttcttttc gcctcacca	420
agggcaaaaa aatccacaat actagcagct aaaacaatac ttgatgcagc cgtaagagt	480
gggtccccgg aaaatataat aggttgata gatgaacctt caattgaact aactcaatat	540
ttaatgcaaa aagcagatat aacccttgca actgggtgct cctcactagt taaatctgct	600
tattcttcgg gaaaaccagc aataggtggt ggtccgggta acacccagc aataattgat	660
gaatctgctc atataaaaat ggcagtaagt tcaattatat tatccaaac ctatgataat	720
gggtgttatat gtgcttctga acaatctgta atagtcttaa aatccatata taacaaggta	780
aaagatgagt tccaagaaag aggagcttat ataataaaga aaacgaatt ggataaagtc	840
cgtgaagtga tttttaaaga tggatcogta aaccctaaaa tagtcggaca gtcagcttat	900
actatagcag ctatggctgg cataaaagta cctaaaacca caagaatatt aataggagaa	960
gttacctcct taggtgaaga agaacctttt gccacgaaa aactatctcc tgttttggt	1020
atgtatgagg ctgacaattt tgatgatgct ttaaaaaaag cagtaactct aataaactta	1080
ggaggcctcg gccatactc aggaatatat gcagatgaaa taaaagcacg agataaaata	1140
gatagattta gtagtgccat gaaaaccgta agaaccttg taaatatccc aacctcaca	1200
ggtgcaagtg gagatctata taattttaga ataccacctt ctttcacgct tggtgcgga	1260
ttttggggag gaaattctgt ttccgagaat gttggtccaa aacatctttt gaatattaaa	1320
accgtagctg aaaggagaga aaacatgctt tggtttagag ttccacataa agtatatttt	1380
aagttcggtt gtcttcaatt tgccttataa gatttaaaag atctaaagaa aaaaagagcc	1440
tttatagtta ctgatagtga cccctataat ttaactatg ttgattcaat aataaaaata	1500
cttgagcacc tagatattga ttttaaagta tttaataagg ttggaagaga agctgatctt	1560
aaaaccataa aaaaagcaac tgaagaaatg tcctccttta tgccagacac tataatagct	1620

-continued

```

ttagtggtga ccctgaaat gagctctgca aagctaattg gggtactata tgaacatcca 1680
gaagtaaaat ttgaagatct tgcaataaaa tttatggaca taagaaagag aatatatact 1740
ttcccaaaac tcggtaaaaa ggctatgtta gttgcaatta caacttctgc tggttccggt 1800
tctgagggtta ctctttttgc tttagtaact gacaataaca ctggaaataa gtacatgtta 1860
gcagattatg aaatgacacc aaatatggca attgtagatg cagaacttat gatgaaaatg 1920
ccaaagggat taaccgctta ttcagggtata gatgcactag taaatagtat agaagcatac 1980
acatccgtat atgcttcaga atacacaaac ggactagcac tagaggcaat acgattaata 2040
tttaaatatt tgcttgaggc ttacaaaaac ggaagaacca atgaaaaagc aagagagaaa 2100
atggctcacg cttcaactat ggtaggtatg gcatccgcta atgcatttct aggtctatgt 2160
cattccatgg caataaaatt aagttcagaa cacaatattc ctagtggcat tgccaatgca 2220
ttactaatag aagaagtaat aaaatttaac gcagttgata atcctgtaaa acaagcccct 2280
tgcccacaat ataagtatcc aaacaccata tttagatatg ctggaattgc agattatata 2340
aagcttgagg gaaatactga tgaggaaaag gtagatctct taattaacaa aatacatgaa 2400
ctaaaaaaag ctttaaatat accaacttca ataaaggatg caggtgtttt ggaggaaaac 2460
ttctattcct cccttgatag aatatctgaa cttgcactag atgatcaatg cacaggcgct 2520
aatcctagat ttctctttac aagttagata aaagaaatgt atataaattg ttttaaaaaa 2580
caaccttaa 2589

```

<210> SEQ ID NO 153

<211> LENGTH: 862

<212> TYPE: PRT

<213> ORGANISM: Clostridium acetobutylicum

<400> SEQUENCE: 153

```

Met Lys Val Thr Thr Val Lys Glu Leu Asp Glu Lys Leu Lys Val Ile
 1             5             10            15
Lys Glu Ala Gln Lys Lys Phe Ser Cys Tyr Ser Gln Glu Met Val Asp
          20             25            30
Glu Ile Phe Arg Asn Ala Ala Met Ala Ala Ile Asp Ala Arg Ile Glu
          35             40            45
Leu Ala Lys Ala Ala Val Leu Glu Thr Gly Met Gly Leu Val Glu Asp
          50             55            60
Lys Val Ile Lys Asn His Phe Ala Gly Glu Tyr Ile Tyr Asn Lys Tyr
          65             70            75            80
Lys Asp Glu Lys Thr Cys Gly Ile Ile Glu Arg Asn Glu Pro Tyr Gly
          85             90            95
Ile Thr Lys Ile Ala Glu Pro Ile Gly Val Val Ala Ala Ile Ile Pro
          100            105            110
Val Thr Asn Pro Thr Ser Thr Thr Ile Phe Lys Ser Leu Ile Ser Leu
          115            120            125
Lys Thr Arg Asn Gly Ile Phe Phe Ser Pro His Pro Arg Ala Lys Lys
          130            135            140
Ser Thr Ile Leu Ala Ala Lys Thr Ile Leu Asp Ala Ala Val Lys Ser
          145            150            155            160
Gly Ala Pro Glu Asn Ile Ile Gly Trp Ile Asp Glu Pro Ser Ile Glu
          165            170            175
Leu Thr Gln Tyr Leu Met Gln Lys Ala Asp Ile Thr Leu Ala Thr Gly
          180            185            190

```

Gly	Pro	Ser	Leu	Val	Lys	Ser	Ala	Tyr	Ser	Ser	Gly	Lys	Pro	Ala	Ile
195						200						205			
Gly	Val	Gly	Pro	Gly	Asn	Thr	Pro	Val	Ile	Ile	Asp	Glu	Ser	Ala	His
210						215						220			
Ile	Lys	Met	Ala	Val	Ser	Ser	Ile	Ile	Leu	Ser	Lys	Thr	Tyr	Asp	Asn
225				230						235				240	
Gly	Val	Ile	Cys	Ala	Ser	Glu	Gln	Ser	Val	Ile	Val	Leu	Lys	Ser	Ile
		245						250						255	
Tyr	Asn	Lys	Val	Lys	Asp	Glu	Phe	Gln	Glu	Arg	Gly	Ala	Tyr	Ile	Ile
		260						265				270			
Lys	Lys	Asn	Glu	Leu	Asp	Lys	Val	Arg	Glu	Val	Ile	Phe	Lys	Asp	Gly
		275				280						285			
Ser	Val	Asn	Pro	Lys	Ile	Val	Gly	Gln	Ser	Ala	Tyr	Thr	Ile	Ala	Ala
290						295				300					
Met	Ala	Gly	Ile	Lys	Val	Pro	Lys	Thr	Thr	Arg	Ile	Leu	Ile	Gly	Glu
305				310						315				320	
Val	Thr	Ser	Leu	Gly	Glu	Glu	Glu	Pro	Phe	Ala	His	Glu	Lys	Leu	Ser
		325						330				335			
Pro	Val	Leu	Ala	Met	Tyr	Glu	Ala	Asp	Asn	Phe	Asp	Asp	Ala	Leu	Lys
		340						345				350			
Lys	Ala	Val	Thr	Leu	Ile	Asn	Leu	Gly	Gly	Leu	Gly	His	Thr	Ser	Gly
		355				360						365			
Ile	Tyr	Ala	Asp	Glu	Ile	Lys	Ala	Arg	Asp	Lys	Ile	Asp	Arg	Phe	Ser
370						375				380					
Ser	Ala	Met	Lys	Thr	Val	Arg	Thr	Phe	Val	Asn	Ile	Pro	Thr	Ser	Gln
385				390						395				400	
Gly	Ala	Ser	Gly	Asp	Leu	Tyr	Asn	Phe	Arg	Ile	Pro	Pro	Ser	Phe	Thr
		405						410						415	
Leu	Gly	Cys	Gly	Phe	Trp	Gly	Gly	Asn	Ser	Val	Ser	Glu	Asn	Val	Gly
		420						425				430			
Pro	Lys	His	Leu	Leu	Asn	Ile	Lys	Thr	Val	Ala	Glu	Arg	Arg	Glu	Asn
		435				440						445			
Met	Leu	Trp	Phe	Arg	Val	Pro	His	Lys	Val	Tyr	Phe	Lys	Phe	Gly	Cys
450						455				460					
Leu	Gln	Phe	Ala	Leu	Lys	Asp	Leu	Lys	Asp	Leu	Lys	Lys	Lys	Arg	Ala
465				470						475				480	
Phe	Ile	Val	Thr	Asp	Ser	Asp	Pro	Tyr	Asn	Leu	Asn	Tyr	Val	Asp	Ser
		485						490						495	
Ile	Ile	Lys	Ile	Leu	Glu	His	Leu	Asp	Ile	Asp	Phe	Lys	Val	Phe	Asn
		500						505				510			
Lys	Val	Gly	Arg	Glu	Ala	Asp	Leu	Lys	Thr	Ile	Lys	Lys	Ala	Thr	Glu
		515				520						525			
Glu	Met	Ser	Ser	Phe	Met	Pro	Asp	Thr	Ile	Ile	Ala	Leu	Gly	Gly	Thr
530						535				540					
Pro	Glu	Met	Ser	Ser	Ala	Lys	Leu	Met	Trp	Val	Leu	Tyr	Glu	His	Pro
545				550						555				560	
Glu	Val	Lys	Phe</												

-continued

Val Thr Asp Asn Asn Thr Gly Asn Lys Tyr Met Leu Ala Asp Tyr Glu	
610 615 620	
Met Thr Pro Asn Met Ala Ile Val Asp Ala Glu Leu Met Met Lys Met	
625 630 635 640	
Pro Lys Gly Leu Thr Ala Tyr Ser Gly Ile Asp Ala Leu Val Asn Ser	
645 650 655	
Ile Glu Ala Tyr Thr Ser Val Tyr Ala Ser Glu Tyr Thr Asn Gly Leu	
660 665 670	
Ala Leu Glu Ala Ile Arg Leu Ile Phe Lys Tyr Leu Pro Glu Ala Tyr	
675 680 685	
Lys Asn Gly Arg Thr Asn Glu Lys Ala Arg Glu Lys Met Ala His Ala	
690 695 700	
Ser Thr Met Ala Gly Met Ala Ser Ala Asn Ala Phe Leu Gly Leu Cys	
705 710 715 720	
His Ser Met Ala Ile Lys Leu Ser Ser Glu His Asn Ile Pro Ser Gly	
725 730 735	
Ile Ala Asn Ala Leu Leu Ile Glu Glu Val Ile Lys Phe Asn Ala Val	
740 745 750	
Asp Asn Pro Val Lys Gln Ala Pro Cys Pro Gln Tyr Lys Tyr Pro Asn	
755 760 765	
Thr Ile Phe Arg Tyr Ala Arg Ile Ala Asp Tyr Ile Lys Leu Gly Gly	
770 775 780	
Asn Thr Asp Glu Glu Lys Val Asp Leu Leu Ile Asn Lys Ile His Glu	
785 790 795 800	
Leu Lys Lys Ala Leu Asn Ile Pro Thr Ser Ile Lys Asp Ala Gly Val	
805 810 815	
Leu Glu Glu Asn Phe Tyr Ser Ser Leu Asp Arg Ile Ser Glu Leu Ala	
820 825 830	
Leu Asp Asp Gln Cys Thr Gly Ala Asn Pro Arg Phe Pro Leu Thr Ser	
835 840 845	
Glu Ile Lys Glu Met Tyr Ile Asn Cys Phe Lys Lys Gln Pro	
850 855 860	

<210> SEQ ID NO 154

<211> LENGTH: 1164

<212> TYPE: DNA

<213> ORGANISM: Escherichia coli

<400> SEQUENCE: 154

atgaacaact ttaatctgca caccccaacc cgcattctgt ttggtaaagg cgcaatcgct	60
ggtttacgcg aacaaattcc tcacgatgct cgcgtattga ttacctacgg cggcggcagc	120
gtgaaaaaaa ccggcggttct cgatcaagtt ctggatgccc tgaaaggcat ggacgtgctg	180
gaatttggcg gtattgagcc aaaccgggct tatgaaacgc tgatgaacgc cgtgaaactg	240
gttcgcgaac agaaagtgc tttcctgctg gcggttggcg gcggttctgt actggacggc	300
accaaattta tcgccgcagc ggctaactat ccggaaaata tcgatccgtg gcacattctg	360
caaacgggag gtaagagat taaaagcgcc atcccgatgg gctgtgtgct gacgctgcc	420
gcaaccgggt cagaatccaa cgcaggcgcg gtgatctccc gtaaaaccac aggcgacaag	480
caggcggtcc attctgccca tgttcagccg gtatttgccg tgctcgatcc ggtttatacc	540
tacaccctgc cgccgcgtca ggtggctaac ggcgtagtgg acgcctttgt acacaccgtg	600
gaacagtatg ttacaaaacc ggttgatgcc aaaattcagg accgtttcgc agaaggcatt	660
ttgctgacgc taatcgaaga tggctcgaaa gccctgaaa agccagaaaa ctacgatgtg	720

-continued

```

cgcgcacaacg tcatgtgggc ggcgactcag gcgtgaaacg gtttgattgg cgctggcgta    780
ccgcaggact gggcaacgca tatgtgggc cacgaactga ctgcgatgca cggctctggat    840
cacgcgcaaa cactggctat cgtcctgcct gcaactgtgga atgaaaaacg cgataccaag    900
cgcgctaagc tgctgcaata tgctgaacgc gtctggaaca tcaactgaagg ttccgatgat    960
gagcgtattg acgccgcgat tgccgcaacc cgcaatttct ttgagcaatt aggcgtgccc    1020
acccacctct ccgactacgg tctggacggc agctccatcc cggctttgct gaaaaaactg    1080
gaagagcacg gcatgaccca actgggcgaa aatcatgaca ttacgttgga tgcagccgc    1140
cgtatatacg aagccgcccg ctaa                                           1164

```

<210> SEQ ID NO 155

<211> LENGTH: 387

<212> TYPE: PRT

<213> ORGANISM: Escherichia coli

<400> SEQUENCE: 155

```

Met Asn Asn Phe Asn Leu His Thr Pro Thr Arg Ile Leu Phe Gly Lys
 1             5             10             15
Gly Ala Ile Ala Gly Leu Arg Glu Gln Ile Pro His Asp Ala Arg Val
          20             25             30
Leu Ile Thr Tyr Gly Gly Gly Ser Val Lys Lys Thr Gly Val Leu Asp
          35             40             45
Gln Val Leu Asp Ala Leu Lys Gly Met Asp Val Leu Glu Phe Gly Gly
          50             55             60
Ile Glu Pro Asn Pro Ala Tyr Glu Thr Leu Met Asn Ala Val Lys Leu
        65             70             75             80
Val Arg Glu Gln Lys Val Thr Phe Leu Leu Ala Val Gly Gly Gly Ser
          85             90             95
Val Leu Asp Gly Thr Lys Phe Ile Ala Ala Ala Ala Asn Tyr Pro Glu
          100            105            110
Asn Ile Asp Pro Trp His Ile Leu Gln Thr Gly Gly Lys Glu Ile Lys
          115            120            125
Ser Ala Ile Pro Met Gly Cys Val Leu Thr Leu Pro Ala Thr Gly Ser
          130            135            140
Glu Ser Asn Ala Gly Ala Val Ile Ser Arg Lys Thr Thr Gly Asp Lys
          145            150            155            160
Gln Ala Phe His Ser Ala His Val Gln Pro Val Phe Ala Val Leu Asp
          165            170            175
Pro Val Tyr Thr Tyr Thr Leu Pro Pro Arg Gln Val Ala Asn Gly Val
          180            185            190
Val Asp Ala Phe Val His Thr Val Glu Gln Tyr Val Thr Lys Pro Val
          195            200            205
Asp Ala Lys Ile Gln Asp Arg Phe Ala Glu Gly Ile Leu Leu Thr Leu
          210            215            220
Ile Glu Asp Gly Pro Lys Ala Leu Lys Glu Pro Glu Asn Tyr Asp Val
          225            230            235            240
Arg Ala Asn Val Met Trp Ala Ala Thr Gln Ala Leu Asn Gly Leu Ile
          245            250            255
Gly Ala Gly Val Pro Gln Asp Trp Ala Thr His Met Leu Gly His Glu
          260            265            270
Leu Thr Ala Met His Gly Leu Asp His Ala Gln Thr Leu Ala Ile Val
          275            280            285

```

-continued

Leu Pro Ala Leu Trp Asn Glu Lys Arg Asp Thr Lys Arg Ala Lys Leu
 290 295 300
 Leu Gln Tyr Ala Glu Arg Val Trp Asn Ile Thr Glu Gly Ser Asp Asp
 305 310 315 320
 Glu Arg Ile Asp Ala Ala Ile Ala Ala Thr Arg Asn Phe Phe Glu Gln
 325 330 335
 Leu Gly Val Pro Thr His Leu Ser Asp Tyr Gly Leu Asp Gly Ser Ser
 340 345 350
 Ile Pro Ala Leu Leu Lys Lys Leu Glu Glu His Gly Met Thr Gln Leu
 355 360 365
 Gly Glu Asn His Asp Ile Thr Leu Asp Val Ser Arg Arg Ile Tyr Glu
 370 375 380
 Ala Ala Arg
 385

<210> SEQ ID NO 156
 <211> LENGTH: 3883
 <212> TYPE: DNA
 <213> ORGANISM: Escherichia coli

<400> SEQUENCE: 156

ctatatgtct gaaggtacag gcgtttccat aactatttgc tcgcgttttt tactcaagaa	60
gaaaatgcc aatagcaaca tcaggcagac aatacccgaa attgcgaaga aaactgtctg	120
gtagcctgcg tgggtcaaaga gtatcccagt cggcgttgaa agcagcacia tccaagcga	180
actggcaatt tgaaaaccaa tcagaaagat cgtcgacgac aggcgcttat caaagtttgc	240
cacgtgtgat ttgaagacgg atatgacaca aagtggaaac tcaatggcat gtaacaactt	300
cactaatgaa ataatccagg ggtaacgaa cagcgcgcag gaaaggatac gcaacgccat	360
aatcacaact ccgataagta atgcattttt tggccctacc cgattcacia agaaaggaat	420
aatcgccatg cacagcgctt cgagtaccac ctggaatgag ttgagataac catacaggcg	480
cgttcctaca tcgtgtgatt cgaataaacc tgaataaaag acaggaaaaa gttgttgatc	540
aaaaatgtta tagaaagacc acgtccccac aataaatatg acgaaaaacc agaagtttcg	600
atccttgaaa actgcgataa aatcctcttt ttttaccctt cccgcatctg ccgctacgca	660
ctggtgatcc ttatctttaa aacgcattgt gatcatcata aatacagcgc caaatagcga	720
gaccaaccag aagttgatat ggggactgat actaaaaaat atgccggcaa agaacgcgcc	780
aatagcatag ccaaaagatc ccaggcgcg cgctgttcca tattcgaaat gaaaatttcg	840
cgccattttt tcggtgaagc tatcaagcaa accgcatccc gccagatacc ccaagccaaa	900
aaatagcgcc ccagaaatta gacctacaga aaaattgctt tgcagtaacg gttcataaac	960
gtaaatcata aacggtccgg tcaagaccag gatgaaactc atacaccaga tgagcggttt	1020
cttcagaccg agtttatcct gaacgatgcc gtgaacatc ataaatagaa tgctggtaaa	1080
ctggttgacc gaataaagtg tacctaattc cgtccctgtc aaccctagat gtcctttcag	1140
ccaaatagcg tataacgacc accacagcga ccaggaaata aaaaagagaa atgagtaact	1200
ggatgcaaaa cgatagtagc cattttctgaa tggaatatcc agtgccataa ttacctgcct	1260
gtcggtaaaa aattcacgtc ctatttagag ataagagcga cttcgccgtt tacttctcac	1320
tattccagtt cttgtcgaca tggcagcgct gtcattgccc ctttcgccgt tactgcaagc	1380
gtcccgcaac gttgagcgag atcgataatt cgtcgcatct ctctctcatc ttagataat	1440
cccgtagagg acagacctgt gagtaaccgg gcaacgaacg catctccgcg ccccggtgcta	1500

-continued

tcgacacaat tcacagacat tccagcaaaa tgggtgaactt gtctctgata acagaccacc	1560
accccttctg cacctttagt caccaacagc atggcgatct catactcttt tgccaggggc	1620
catatatcct gatcggtctg tgtttttcca ctgataagtc gccattcttc ttcgagagc	1680
ttgacgacat cgcgcagttg tagcgctgc cgcaaacaca agcggagcaa atgctcgtct	1740
tgccatagat cttcacgaat attaggatcg aagctgacaa aacctccggc atgccggatc	1800
gccgtcatcg cagtaaatgc gctggtacgc gaaggctcgg cagacaacgc aattgaacag	1860
agatgtaacc attcgccatg tcgccagcag ggcaagctcg tcgtctctaa aaaaagatcg	1920
gcactggggc ggaccataaa cgtaaatgaa cgttccccct gatcggttcag atcgacaagc	1980
accgtggatg tccggtgcc ttcattctgc ttcagatacg tgatatcgac tccctcagtt	2040
agcagcgctt tttgcattaa cgcaccaaaa ggatcatccc ccaccgacc tataaaccca	2100
cttggtccgc ctaatctggc gattcccccc gcaacgtag ctggcgcgcc gccaggacaa	2160
ggcagtaggc gcccgctga tcttggaag agatctacga ccgcacccc taaaacccat	2220
actttggctg acattttttt cccttaaat catctgagtt acgcatagtg ataaacctct	2280
ttttcgaaa atcgctcatg atttactaaa acatgcatat tcgatcacia aacgtcatag	2340
ttaacgttaa catttgtgat attcatcgca tttatgaaag taagggactt tatttttata	2400
aaagttaacg ttaacaattc accaaatttg cttaaccagg atgattaaaa tgacgcaatc	2460
tcgattgcat gcggcgcaaa acgcccagc aaaacttcat gagcacggg gtaacacttt	2520
ctatccccat tttcacctcg cgcctcctgc cgggtggatg aacgatccaa acggcctgat	2580
ctggtttaac gatcgttatc acgcgtttta tcaacatcat ccgatgagcg aacactgggg	2640
gccaatgcac tggggacatg ccaccagcga cgatatgac cactggcagc atgagcctat	2700
tgcgctagcg ccaggagacg ataatgacaa agacgggtgt ttttcaggta gtgctgtcga	2760
tgacaatggt gtcctctcac ttatctacac cggacacgtc tggctcgatg gtgcaggtaa	2820
tgacgatgca attcgcaag tacaatgtct ggctaccagt cgggatggta ttcatttcga	2880
gaaacagggt gtgatcctca ctccaccaga aggaatcatg cacttccgag atcctaaagt	2940
gtggcgtgaa gccgacacat ggtggatggt agtcggggcg aaagatccag gcaacacggg	3000
gcagatcctg ctttatcgcg gcagttcgtt gcgtgaatgg accttcgac gcgtactggc	3060
ccacgctgat gcgggtgaaa gctatatgtg ggaatgtccg gactttttca gccttggcga	3120
tcagcattat ctgatgtttt ccccgaggg aatgaatgcc gagggataca gttaccgaaa	3180
tcgctttcaa agtggcgtaa tacccggaat gtggtcgcca ggacgacttt ttgcacaatc	3240
cgggcatttt actgaacttg ataacgggca tgacttttat gcaccacaaa gcttttttagc	3300
gaaggatggt cggcgatattg ttatcggtg gatggatatg tgggaatcgc caatgcctc	3360
aaaacgtgaa ggatgggcag gctgcatgac gctggcgcg gagctatcag agagcaatgg	3420
caaatctcta caacgcccgg tacacgaagc tgagtcgtta cgccagcagc atcaatctgt	3480
ctctccccgc acaatcagca ataaatatgt tttgcaggaa aacgcgcaag cagttgagat	3540
tcagttgcag tgggcgctga agaacagtga tgccgaacat tacggattac agctcggcac	3600
tggaatgcgg ctgtatatg ataaccaatc tgagcgactt gttttgtggc ggtattacc	3660
acacgagaat ttagacggct accgtagtat tcccccccc cagcgtgaca cgctcgcct	3720
aaggatattt atcgatacat catccgtgga agtatattat aacgacgggg aagcgggtgat	3780
gagtagtcga atctatccgc agccagaaga acgggaactg tcgctttatg cctcccacgg	3840
agtggctgtg ctgcaacatg gagcactctg gctactgggt taa	3883

-continued

```

<210> SEQ ID NO 157
<211> LENGTH: 477
<212> TYPE: PRT
<213> ORGANISM: Escherichia coli

<400> SEQUENCE: 157

Met Thr Gln Ser Arg Leu His Ala Ala Gln Asn Ala Leu Ala Lys Leu
1          5          10          15

His Glu His Arg Gly Asn Thr Phe Tyr Pro His Phe His Leu Ala Pro
20          25          30

Pro Ala Gly Trp Met Asn Asp Pro Asn Gly Leu Ile Trp Phe Asn Asp
35          40          45

Arg Tyr His Ala Phe Tyr Gln His His Pro Met Ser Glu His Trp Gly
50          55          60

Pro Met His Trp Gly His Ala Thr Ser Asp Asp Met Ile His Trp Gln
65          70          75          80

His Glu Pro Ile Ala Leu Ala Pro Gly Asp Asp Asn Asp Lys Asp Gly
85          90          95

Cys Phe Ser Gly Ser Ala Val Asp Asp Asn Gly Val Leu Ser Leu Ile
100         105         110

Tyr Thr Gly His Val Trp Leu Asp Gly Ala Gly Asn Asp Asp Ala Ile
115         120         125

Arg Glu Val Gln Cys Leu Ala Thr Ser Arg Asp Gly Ile His Phe Glu
130         135         140

Lys Gln Gly Val Ile Leu Thr Pro Pro Glu Gly Ile Met His Phe Arg
145         150         155         160

Asp Pro Lys Val Trp Arg Glu Ala Asp Thr Trp Trp Met Val Val Gly
165         170         175

Ala Lys Asp Pro Gly Asn Thr Gly Gln Ile Leu Leu Tyr Arg Gly Ser
180         185         190

Ser Leu Arg Glu Trp Thr Phe Asp Arg Val Leu Ala His Ala Asp Ala
195         200         205

Gly Glu Ser Tyr Met Trp Glu Cys Pro Asp Phe Phe Ser Leu Gly Asp
210         215         220

Gln His Tyr Leu Met Phe Ser Pro Gln Gly Met Asn Ala Glu Gly Tyr
225         230         235         240

Ser Tyr Arg Asn Arg Phe Gln Ser Gly Val Ile Pro Gly Met Trp Ser
245         250         255

Pro Gly Arg Leu Phe Ala Gln Ser Gly His Phe Thr Glu Leu Asp Asn
260         265         270

Gly His Asp Phe Tyr Ala Pro Gln Ser Phe Leu Ala Lys Asp Gly Arg
275         280         285

Arg Ile Val Ile Gly Trp Met Asp Met Trp Glu Ser Pro Met Pro Ser
290         295         300

Lys Arg Glu Gly Trp Ala Gly Cys Met Thr Leu Ala Arg Glu Leu Ser
305         310         315         320

Glu Ser Asn Gly Lys Leu Leu Gln Arg Pro Val His Glu Ala Glu Ser
325         330         335

Leu Arg Gln Gln His Gln Ser Val Ser Pro Arg Thr Ile Ser Asn Lys
340         345         350

Tyr Val Leu Gln Glu Asn Ala Gln Ala Val Glu Ile Gln Leu Gln Trp
355         360         365

```

-continued

Ala	Leu	Lys	Asn	Ser	Asp	Ala	Glu	His	Tyr	Gly	Leu	Gln	Leu	Gly	Thr
370						375					380				
Gly	Met	Arg	Leu	Tyr	Ile	Asp	Asn	Gln	Ser	Glu	Arg	Leu	Val	Leu	Trp
385					390					395					400
Arg	Tyr	Tyr	Pro	His	Glu	Asn	Leu	Asp	Gly	Tyr	Arg	Ser	Ile	Pro	Leu
				405					410					415	
Pro	Gln	Arg	Asp	Thr	Leu	Ala	Leu	Arg	Ile	Phe	Ile	Asp	Thr	Ser	Ser
			420					425					430		
Val	Glu	Val	Phe	Ile	Asn	Asp	Gly	Glu	Ala	Val	Met	Ser	Ser	Arg	Ile
	435						440					445			
Tyr	Pro	Gln	Pro	Glu	Glu	Arg	Glu	Leu	Ser	Leu	Tyr	Ala	Ser	His	Gly
	450					455					460				
Val	Ala	Val	Leu	Gln	His	Gly	Ala	Leu	Trp	Leu	Leu	Gly			
465					470					475					

<210> SEQ ID NO 158
 <211> LENGTH: 304
 <212> TYPE: PRT
 <213> ORGANISM: Escherichia coli

<400> SEQUENCE: 158

Met	Ser	Ala	Lys	Val	Trp	Val	Leu	Gly	Asp	Ala	Val	Val	Asp	Leu	Leu
1				5					10					15	
Pro	Glu	Ser	Asp	Gly	Arg	Leu	Leu	Pro	Cys	Pro	Gly	Gly	Ala	Pro	Ala
			20					25					30		
Asn	Val	Ala	Val	Gly	Ile	Ala	Arg	Leu	Gly	Gly	Thr	Ser	Gly	Phe	Ile
	35					40						45			
Gly	Arg	Val	Gly	Asp	Asp	Pro	Phe	Gly	Ala	Leu	Met	Gln	Arg	Thr	Leu
	50				55						60				
Leu	Thr	Glu	Gly	Val	Asp	Ile	Thr	Tyr	Leu	Lys	Gln	Asp	Glu	Trp	His
65				70					75					80	
Arg	Thr	Ser	Thr	Val	Leu	Val	Asp	Leu	Asn	Asp	Gln	Gly	Glu	Arg	Ser
			85					90						95	
Phe	Thr	Phe	Met	Val	Arg	Pro	Ser	Ala	Asp	Leu	Phe	Leu	Glu	Thr	Thr
			100					105					110		
Asp	Leu	Pro	Cys	Trp	Arg	His	Gly	Glu	Trp	Leu	His	Leu	Cys	Ser	Ile
	115					120						125			
Ala	Leu	Ser	Ala	Glu	Pro	Ser	Arg	Thr	Ser	Ala	Phe	Thr	Ala	Met	Thr
	130					135						140			
Ala	Ile	Arg	His	Ala	Gly	Gly	Phe	Val	Ser	Phe	Asp	Pro	Asn	Ile	Arg
145				150						155				160	
Glu	Asp	Leu	Trp	Gln	Asp	Glu	His	Leu	Leu	Arg	Leu	Cys	Leu	Arg	Gln
			165					170						175	
Ala	Leu	Gln	Leu	Ala	Asp	Val	Val	Lys	Leu	Ser	Glu	Glu	Glu	Trp	Arg
			180					185						190	
Leu	Ile	Ser	Gly	Lys	Thr	Gln	Asn	Asp	Gln	Asp	Ile	Cys	Ala	Leu	Ala
	195					200						205			
Lys	Glu	Tyr	Glu	Ile	Ala	Met	Leu	Leu	Val	Thr	Lys	Gly	Ala	Glu	Gly
	210					215					220				
Val	Val	Val	Cys	Tyr	Arg	Gly	Gln	Val	His	His	Phe	Ala	Gly	Met	Ser
225					230					235				240	
Val	Asn	Cys	Val	Asp	Ser	Thr	Gly	Ala	Gly	Asp	Ala	Phe	Val	Ala	Gly
			245					250						255	
Leu	Leu	Thr	Gly	Leu	Ser	Ser	Thr	Gly	Leu	Ser	Thr	Asp	Glu	Arg	Glu
			260					265						270	

-continued

Met Arg Arg Ile Ile Asp Leu Ala Gln Arg Cys Gly Ala Leu Ala Val
275 280 285

Thr Ala Lys Gly Ala Met Thr Ala Leu Pro Cys Arg Gln Glu Leu Glu
290 295 300

<210> SEQ ID NO 159
<211> LENGTH: 415
<212> TYPE: PRT
<213> ORGANISM: Escherichia coli

<400> SEQUENCE: 159

Met Ala Leu Asn Ile Pro Phe Arg Asn Ala Tyr Tyr Arg Phe Ala Ser
1 5 10 15

Ser Tyr Ser Phe Leu Phe Phe Ile Ser Trp Ser Leu Trp Trp Ser Leu
20 25 30

Tyr Ala Ile Trp Leu Lys Gly His Leu Gly Leu Thr Gly Thr Glu Leu
35 40 45

Gly Thr Leu Tyr Ser Val Asn Gln Phe Thr Ser Ile Leu Phe Met Met
50 55 60

Phe Tyr Gly Ile Val Gln Asp Lys Leu Gly Leu Lys Lys Pro Leu Ile
65 70 75 80

Trp Cys Met Ser Phe Ile Leu Val Leu Thr Gly Pro Phe Met Ile Tyr
85 90 95

Val Tyr Glu Pro Leu Leu Gln Ser Asn Phe Ser Val Gly Leu Ile Leu
100 105 110

Gly Ala Leu Phe Phe Gly Leu Gly Tyr Leu Ala Gly Cys Gly Leu Leu
115 120 125

Asp Ser Phe Thr Glu Lys Met Ala Arg Asn Phe His Phe Glu Tyr Gly
130 135 140

Thr Ala Arg Ala Trp Gly Ser Phe Gly Tyr Ala Ile Gly Ala Phe Phe
145 150 155 160

Ala Gly Ile Phe Phe Ser Ile Ser Pro His Ile Asn Phe Trp Leu Val
165 170 175

Ser Leu Phe Gly Ala Val Phe Met Met Ile Asn Met Arg Phe Lys Asp
180 185 190

Lys Asp His Gln Cys Val Ala Ala Asp Ala Gly Gly Val Lys Lys Glu
195 200 205

Asp Phe Ile Ala Val Phe Lys Asp Arg Asn Phe Trp Val Phe Val Ile
210 215 220

Phe Ile Val Gly Thr Trp Ser Phe Tyr Asn Ile Phe Asp Gln Gln Leu
225 230 235 240

Phe Pro Val Phe Tyr Ser Gly Leu Phe Glu Ser His Asp Val Gly Thr
245 250 255

Arg Leu Tyr Gly Tyr Leu Asn Ser Phe Gln Val Val Leu Glu Ala Leu
260 265 270

Cys Met Ala Ile Ile Pro Phe Phe Val Asn Arg Val Gly Pro Lys Asn
275 280 285

Ala Leu Leu Ile Gly Val Val Ile Met Ala Leu Arg Ile Leu Ser Cys
290 295 300

Ala Leu Phe Val Asn Pro Trp Ile Ile Ser Leu Val Lys Leu Leu His
305 310 315 320

Ala Ile Glu Val Pro Leu Cys Val Ile Ser Val Phe Lys Tyr Ser Val
325 330 335

-continued

Ala Asn Phe Asp Lys Arg Leu Ser Ser Thr Ile Phe Leu Ile Gly Phe
 340 345 350

Gln Ile Ala Ser Ser Leu Gly Ile Val Leu Leu Ser Thr Pro Thr Gly
 355 360 365

Ile Leu Phe Asp His Ala Gly Tyr Gln Thr Val Phe Phe Ala Ile Ser
 370 375 380

Gly Ile Val Cys Leu Met Leu Leu Phe Gly Ile Phe Phe Leu Ser Lys
 385 390 395 400

Lys Arg Glu Gln Ile Val Met Glu Thr Pro Val Pro Ser Ala Ile
 405 410 415

<210> SEQ ID NO 160
 <211> LENGTH: 18
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Primer Scr1

<400> SEQUENCE: 160

cctttcttttg tgaatcgg

18

<210> SEQ ID NO 161
 <211> LENGTH: 18
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Primer Scr2

<400> SEQUENCE: 161

agaaacaggg tgtgatcc

18

<210> SEQ ID NO 162
 <211> LENGTH: 20
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Primer Scr3

<400> SEQUENCE: 162

agtgatcatc acctgttgcc

20

<210> SEQ ID NO 163
 <211> LENGTH: 20
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Primer Scr4

<400> SEQUENCE: 163

agcacggcga ggtcgacgg

20

<210> SEQ ID NO 164
 <211> LENGTH: 74
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Primer OT731

<400> SEQUENCE: 164

aaagctggag ctccaccgcg gtggcggcgg ctctagaagt tttcaaagca gagtttcggt

60

tgaatatattt acca

74

<210> SEQ ID NO 165

-continued

<211> LENGTH: 50
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Primer OT732

 <400> SEQUENCE: 165

 ttcaatatgc atgcctcaga acgtttacat tgtatcgact gccagaaccc 50

 <210> SEQ ID NO 166
 <211> LENGTH: 79
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Primer OT733

 <400> SEQUENCE: 166

 gcagtcgata caatgtaaac gttctgaggc atgcatattg aattttcaaa aattcttact 60
 ttttttttgg atggacgca 79

 <210> SEQ ID NO 167
 <211> LENGTH: 72
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Primer OT734

 <400> SEQUENCE: 167

 acctgcacct ataacacata ctttttccat ggtagttttt tctccttgac gttaaagtat 60
 agaggtatat ta 72

 <210> SEQ ID NO 168
 <211> LENGTH: 60
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Primer OT735

 <400> SEQUENCE: 168

 aaaaactacc atggaagagg tatgtgttat aggtgcaggt actatgggtt caggaattgc 60

 <210> SEQ ID NO 169
 <211> LENGTH: 71
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Primer OT736

 <400> SEQUENCE: 169

 gtaaaaaaaaa gaaggccgta taggccttat tttgaataat cgtagaaacc ttttcctgat 60
 tttcttccaa g 71

 <210> SEQ ID NO 170
 <211> LENGTH: 81
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Primer OT737

 <400> SEQUENCE: 170

 acgattattc aaaataaggc ctatacggcc ttcttttttt tactttgttc agaacaactt 60
 ctcatttttt tctactcata a 81

-continued

<210> SEQ ID NO 171
 <211> LENGTH: 73
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Primer OT738

 <400> SEQUENCE: 171

 gaattgggta cggggccccc cctcgaggtc gaccgatgcc tcataaactt cggtagttat 60
 attactctga gat 73

 <210> SEQ ID NO 172
 <211> LENGTH: 65
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Primer OT797

 <400> SEQUENCE: 172

 aaagtaagaa tttttgaaaa ttcaatatgc atgcaagaag ttgtaatagc tagtgagta 60
 agaac 65

 <210> SEQ ID NO 173
 <211> LENGTH: 73
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Primer OT798

 <400> SEQUENCE: 173

 gaaaaagatc atgagaaaat cgcagaacgt aaggcgcgcc tcagcacttt tctagcaata 60
 ttgctgttcc ttg 73

 <210> SEQ ID NO 174
 <211> LENGTH: 41
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Primer OT806

 <400> SEQUENCE: 174

 ctcgaaaata gggcgcgcc ccattaccga catttggcg c 41

 <210> SEQ ID NO 175
 <211> LENGTH: 55
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Primer OT807

 <400> SEQUENCE: 175

 actgcactag ctattacaac ttcttgcatt cgtgatgatt gattgattga ttgta 55

 <210> SEQ ID NO 176
 <211> LENGTH: 55
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Primer OT808

 <400> SEQUENCE: 176

 actgcactag ctattacaac ttcttgcatt cgtgatgatt gattgattga ttgta 55

 <210> SEQ ID NO 177

-continued

<211> LENGTH: 60
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer OT809

<400> SEQUENCE: 177

tttcgaataa acacacataa acaaacaccc catggaaaag gtatgtgtta taggtgcagg 60

<210> SEQ ID NO 178
<211> LENGTH: 62
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer OT799

<400> SEQUENCE: 178

taccggggccc cccctcgagg tcgacggcgc gccactggta gagagcgact ttgtatgccc 60

ca 62

<210> SEQ ID NO 179
<211> LENGTH: 60
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer OT761

<400> SEQUENCE: 179

cttggccttc actagcatgc tgaatatgta ttacttgggt atggttatat atgacaaaag 60

<210> SEQ ID NO 180
<211> LENGTH: 68
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer OT803

<400> SEQUENCE: 180

ccctcactaa aggggaacaaa agctggagct cgatatcggc gcgccacat gcagtgatgc 60

acgcgcga 68

<210> SEQ ID NO 181
<211> LENGTH: 49
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer OT804

<400> SEQUENCE: 181

aaggatgaca ttgtttagtt ccatggttgt aatatgtgtg tttgtttgg 49

<210> SEQ ID NO 182
<211> LENGTH: 51
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer OT785

<400> SEQUENCE: 182

cacacatatt acaaccatgg aactaaacaa tgatcatcctt gaaaaggaag g 51

<210> SEQ ID NO 183
<211> LENGTH: 63
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence

-continued

<220> FEATURE:
 <223> OTHER INFORMATION: Primer OT786

<400> SEQUENCE: 183

atcattcatt ggccattcag gccttatcta tttttgaagc cttcaatttt tcttttctct	60
atg	63

<210> SEQ ID NO 184
 <211> LENGTH: 65
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Primer OT787

<400> SEQUENCE: 184

caaaaataga taaggcctga atggccaatg aatgatttga tgattttctt ttcctccat	60
ttttc	65

<210> SEQ ID NO 185
 <211> LENGTH: 77
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Primer PT805

<400> SEQUENCE: 185

gaattgggta cggggccccc cctcgaggtc gacttatagt attatatatt ctgatttgg	60
tatagcaagc agcgttt	77

<210> SEQ ID NO 186
 <211> LENGTH: 1269
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Codon optimized ter

<400> SEQUENCE: 186

actagtacca taaccaagta atacatattc agcatgctag tgaaggccaa gtctgtcaag	60
ggctttatta gagatgttca tccgtatggg tgtaggagag aagtgttaaa ccagattgac	120
tactgcaaga aagcaattgg ctttaggggc cctaaaaagg ttcttattgt aggtgcttct	180
tcaggcttcg gactagctac tagaatatct gttgcattcg gagggcctga agccataca	240
atcgtgtgtt catacgagac tggagctaca gacagaagga taggtacggc tgggtggtag	300
aataatatct tctttaaaga attcgctaag aagaaagggt tgggtggcaa gaatttcata	360
gaagatgcat tttcgaatga aaccaaggat aaagtataaa agtatataaa ggacgaattt	420
ggtaaaattg atttattcgt atattcttta gctgctccta gaagaaagga ctacaaaacc	480
ggtaatgttt atacctcaag aattaaaaca attctaggtg actttgaagg gcctactatt	540
gacgtagaaa gagatgaat aactttaaag aaggtatctt ctgctagtat cgaggaaatc	600
gaagaacac gtaaagtaat gggcggagaa gactggcagg agtgggtgtga ggagttatta	660
tacgaagatt gtttttctga taaagctaca accatcgctt attcctatat tggcagtcct	720
agaacttata aaatatatcg tgaaggaacc attgggattg ctaagaagga tttagaagac	780
aaagccaagt tgatcaacga aaagcttaat agagtcatag gaggtagggc atttgtgtct	840
gttaacaaag ctttagtaac caaggcatct gcttatattc caaccttccc tctatacgct	900
gccatattat ataaagtaat gaaagaaaag aacattcacg aaaattgtat tatgcaaatt	960

-continued

```

gagcgtatgt tctcagagaa aatatactcc aacgaaaaga ttcagttcga tgataagggc 1020
cgtcttagaa tggacgattt agaactaaga aaggatgttc aggatgaagt tgacagaatt 1080
tggcttaaca taacaccaga aaacttcaag gagcttagtg actacaaggg gtataagaaa 1140
gagtttatga atctaaatgg ttttgattta gatggagttg attattccaa ggatcttgat 1200
attgaattac ttagaaaact agagccttaa gcggccgcgt taattcaaat taattgatat 1260
agtactagt                                     1269

```

```

<210> SEQ ID NO 187
<211> LENGTH: 398
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Modified Ter protein

```

```

<400> SEQUENCE: 187

```

```

Met Leu Val Lys Ala Lys Phe Val Lys Gly Phe Ile Arg Asp Val His
1           5           10          15
Pro Tyr Gly Cys Arg Arg Glu Val Leu Asn Gln Ile Asp Tyr Cys Lys
          20          25          30
Lys Ala Ile Gly Phe Arg Gly Pro Lys Lys Val Leu Ile Val Gly Ala
          35          40          45
Ser Ser Gly Phe Gly Leu Ala Thr Arg Ile Ser Val Ala Phe Gly Gly
          50          55          60
Pro Glu Ala His Thr Ile Gly Val Ser Tyr Glu Thr Gly Ala Thr Asp
65          70          75          80
Arg Arg Ile Gly Thr Ala Gly Trp Tyr Asn Asn Ile Phe Phe Lys Glu
          85          90          95
Phe Ala Lys Lys Lys Gly Leu Val Ala Lys Asn Phe Ile Glu Asp Ala
          100         105         110
Phe Ser Asn Glu Thr Lys Asp Lys Val Ile Lys Tyr Ile Lys Asp Glu
          115         120         125
Phe Gly Lys Ile Asp Leu Phe Val Tyr Ser Leu Ala Ala Pro Arg Arg
          130         135         140
Lys Asp Tyr Lys Thr Gly Asn Val Tyr Thr Ser Arg Ile Lys Thr Ile
          145         150         155         160
Leu Gly Asp Phe Glu Gly Pro Thr Ile Asp Val Glu Arg Asp Glu Ile
          165         170         175
Thr Leu Lys Lys Val Ser Ser Ala Ser Ile Glu Glu Ile Glu Glu Thr
          180         185         190
Arg Lys Val Met Gly Gly Glu Asp Trp Gln Glu Trp Cys Glu Glu Leu
          195         200         205
Leu Tyr Glu Asp Cys Phe Ser Asp Lys Ala Thr Thr Ile Ala Tyr Ser
          210         215         220
Tyr Ile Gly Ser Pro Arg Thr Tyr Lys Ile Tyr Arg Glu Gly Thr Ile
          225         230         235         240
Gly Ile Ala Lys Lys Asp Leu Glu Asp Lys Ala Lys Leu Ile Asn Glu
          245         250         255
Lys Leu Asn Arg Val Ile Gly Gly Arg Ala Phe Val Ser Val Asn Lys
          260         265         270
Ala Leu Val Thr Lys Ala Ser Ala Tyr Ile Pro Thr Phe Pro Leu Tyr
          275         280         285
Ala Ala Ile Leu Tyr Lys Val Met Lys Glu Lys Asn Ile His Glu Asn
          290         295         300

```

-continued

Cys Ile Met Gln Ile Glu Arg Met Phe Ser Glu Lys Ile Tyr Ser Asn
 305 310 315 320
 Glu Lys Ile Gln Phe Asp Asp Lys Gly Arg Leu Arg Met Asp Asp Leu
 325 330 335
 Glu Leu Arg Lys Asp Val Gln Asp Glu Val Asp Arg Ile Trp Ser Asn
 340 345 350
 Ile Thr Pro Glu Asn Phe Lys Glu Leu Ser Asp Tyr Lys Gly Tyr Lys
 355 360 365
 Lys Glu Phe Met Asn Leu Asn Gly Phe Asp Leu Asp Gly Val Asp Tyr
 370 375 380
 Ser Lys Asp Leu Asp Ile Glu Leu Leu Arg Lys Leu Glu Pro
 385 390 395

<210> SEQ ID NO 188
 <211> LENGTH: 1484
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Codon optimized ald

<400> SEQUENCE: 188

actagttcga ataacacac ataacaaac accatggata aagatacctt aatcccaacc	60
accaaagact tgaaagtga gactaatggt gaaaacatca acttaaagaa ttacaaagat	120
aactcttcat gttttggagt atttgaatg gttgagaatg ccatttcttc tgcagtacat	180
gcacaaaaga ttctttccct acactacaca aggaacaaa gagagaaaat aatcaccgaa	240
ataagaaaag ccgcattaca gaataaagag gtcttagcca caatgatcct ggaggaaacc	300
cacatgggaa ggtatgagga taaatcttg aaacatgaat tagtggccaa gtatacccca	360
ggcactgaag atctgacaac aacagcatgg tccggcgata atggactaac agtgggtgaa	420
atgagtccat acggagttaa cggcgctata actccaagca cgaatccaac agaaaccgtt	480
atctgcaatt ctataggat gatagctgcg gggaatgcag ttgtatttaa tggtcaccca	540
tgcgccccaa agtgtgtgcg ttctgcagta gaaatgataa acaaagccat aattagctgt	600
ggtggacctg aaaccttgt cactactata aagaacccaa ctatggaaa tttagacgct	660
attatcaaac atccatccat aaaattgttg tgcggtacgg gtggcccggt tatggtaaaa	720
acccttctta attctggtta aaaggccatc ggagctggcg cgggtaatcc tccggttatt	780
gtagacgata cagcagatat cgagaaggcc ggcagaagca ttattgaagg ttgttcgttt	840
gacaacaatc ttccttgat cgcggaaaaa gaagtgttcg tgtttgaata cgttcagat	900
gatctgatct ctaacatggt gaaaaacaac gccgtcatta tcaatgaaga ccaagtatcc	960
aagctgatag acctgtgtct tcaaaagaac aatgaaactc aagaatattt cattaataag	1020
aagtgggttg gtaaggacgc taaactgttt ttggatgaaa tagatgtaga gtcaccaagt	1080
aatgtaaagt gtattatttg tgaagtcaac gcaaaccatc cgttcgttat gacggagttg	1140
atgatgcaa ttttgcctat agttagagtg aaggacattg atgaagccat taaatacgcc	1200
aagatagctg agcagaatag aaacattcc gcctacattt attctaagaa catcgataac	1260
cttaatagat tcgaacgtga aattgataca actatctttg ttaagaatgc aaagtcattt	1320
gcaggtgtcg gttatgaagc tgagggtttc acaaccttta caattgccgg atccacaggt	1380
gaaggaatca cgtcagctag aaactttacc aggcaaagac gttgtgtcct agcagggttag	1440
ggcctgcagg gccgtgaatt tactttaaat cttgcattac tagt	1484

-continued

```

<210> SEQ ID NO 189
<211> LENGTH: 468
<212> TYPE: PRT
<213> ORGANISM: Artificial
<220> FEATURE:
<223> OTHER INFORMATION: Modified Ald

<400> SEQUENCE: 189

Met Asp Lys Asp Thr Leu Ile Pro Thr Thr Lys Asp Leu Lys Val Lys
1          5          10          15

Thr Asn Gly Glu Asn Ile Asn Leu Lys Asn Tyr Lys Asp Asn Ser Ser
20          25          30

Cys Phe Gly Val Phe Glu Asn Val Glu Asn Ala Ile Ser Ser Ala Val
35          40          45

His Ala Gln Lys Ile Leu Ser Leu His Tyr Thr Lys Glu Gln Arg Glu
50          55          60

Lys Ile Ile Thr Glu Ile Arg Lys Ala Ala Leu Gln Asn Lys Glu Val
65          70          75          80

Leu Ala Thr Met Ile Leu Glu Glu Thr His Met Gly Arg Tyr Glu Asp
85          90          95

Lys Ile Leu Lys His Glu Leu Val Ala Lys Tyr Thr Pro Gly Thr Glu
100         105         110

Asp Leu Thr Thr Thr Ala Trp Ser Gly Asp Asn Gly Leu Thr Val Val
115         120         125

Glu Met Ser Pro Tyr Gly Val Ile Gly Ala Ile Thr Pro Ser Thr Asn
130         135         140

Pro Thr Glu Thr Val Ile Cys Asn Ser Ile Gly Met Ile Ala Ala Gly
145         150         155         160

Asn Ala Val Val Phe Asn Gly His Pro Cys Ala Lys Lys Cys Val Ala
165         170         175

Phe Ala Val Glu Met Ile Asn Lys Ala Ile Ile Ser Cys Gly Gly Pro
180         185         190

Glu Asn Leu Val Thr Thr Ile Lys Asn Pro Thr Met Glu Ser Leu Asp
195         200         205

Ala Ile Ile Lys His Pro Ser Ile Lys Leu Leu Cys Gly Thr Gly Gly
210         215         220

Pro Gly Met Val Lys Thr Leu Leu Asn Ser Gly Lys Lys Ala Ile Gly
225         230         235         240

Ala Gly Ala Gly Asn Pro Pro Val Ile Val Asp Asp Thr Ala Asp Ile
245         250         255

Glu Lys Ala Gly Arg Ser Ile Ile Glu Gly Cys Ser Phe Asp Asn Asn
260         265         270

Leu Pro Cys Ile Ala Glu Lys Glu Val Phe Val Phe Glu Asn Val Ala
275         280         285

Asp Asp Leu Ile Ser Asn Met Leu Lys Asn Asn Ala Val Ile Ile Asn
290         295         300

Glu Asp Gln Val Ser Lys Leu Ile Asp Leu Val Leu Gln Lys Asn Asn
305         310         315         320

Glu Thr Gln Glu Tyr Phe Ile Asn Lys Lys Trp Val Gly Lys Asp Ala
325         330         335

Lys Leu Phe Leu Asp Glu Ile Asp Val Glu Ser Pro Ser Asn Val Lys
340         345         350

Cys Ile Ile Cys Glu Val Asn Ala Asn His Pro Phe Val Met Thr Glu
355         360         365

```

-continued

Leu Met Met Pro Ile Leu Pro Ile Val Arg Val Lys Asp Ile Asp Glu
 370 375 380

Ala Ile Lys Tyr Ala Lys Ile Ala Glu Gln Asn Arg Lys His Ser Ala
 385 390 395 400

Tyr Ile Tyr Ser Lys Asn Ile Asp Asn Leu Asn Arg Phe Glu Arg Glu
 405 410 415

Ile Asp Thr Thr Ile Phe Val Lys Asn Ala Lys Ser Phe Ala Gly Val
 420 425 430

Gly Tyr Glu Ala Glu Gly Phe Thr Thr Phe Thr Ile Ala Gly Ser Thr
 435 440 445

Gly Glu Gly Ile Thr Ser Ala Arg Asn Phe Thr Arg Gln Arg Arg Cys
 450 455 460

Val Leu Ala Gly
 465

<210> SEQ ID NO 190
 <211> LENGTH: 69
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Primer PT800

<400> SEQUENCE: 190

gggaacaaaa gctggagctc caccgcggtg gggcgcgccc tattttcgag gaccttgta 60
 ccttgagcc 69

<210> SEQ ID NO 191
 <211> LENGTH: 50
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Primer OT758

<400> SEQUENCE: 191

ttaaggatc tttatccatg gtgtttgttt atgtgtgttt attcgaaact 50

<210> SEQ ID NO 192
 <211> LENGTH: 52
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Primer OT754

<400> SEQUENCE: 192

ttgggtaccg ggccccccct cgaggctgac tggccattaa tctttcccat at 52

<210> SEQ ID NO 193
 <211> LENGTH: 52
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Primer OT755

<400> SEQUENCE: 193

tgtgtcctag caggttaggg cctgcagggc cgtgaattta ctttaaactc tg 52

<210> SEQ ID NO 194
 <211> LENGTH: 50
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Primer OT760

-continued

<400> SEQUENCE: 194

cgaaaaatagg gcgcgccact ggtagagagc gactttgtat gcccgaattg 50

<210> SEQ ID NO 195

<211> LENGTH: 71

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Primer OT792

<400> SEQUENCE: 195

cccttgacga acttggcctt cactagcatg ctgaatatgt attacttggt tatggttata 60

tatgacaaaa g 71

<210> SEQ ID NO 196

<211> LENGTH: 71

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Primer OT791

<400> SEQUENCE: 196

cccttgacga acttggcctt cactagcatg ctgaatatgt attacttggt tatggttata 60

tatgacaaaa g 71

<210> SEQ ID NO 197

<211> LENGTH: 73

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Primer OT765

<400> SEQUENCE: 197

ggaacaaaag ctggagctcc accgcggtgg tttaacgtat agacttctaa tatatttctc 60

catacttggt att 73

<210> SEQ ID NO 198

<211> LENGTH: 29

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Primer LDHEcoRV F

<400> SEQUENCE: 198

gacgtcatga ccaccgcgcg atccctttt 29

<210> SEQ ID NO 199

<211> LENGTH: 30

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Primer LDH Aat1R

<400> SEQUENCE: 199

gatatccaac accagcgacc gacgtattac 30

<210> SEQ ID NO 200

<211> LENGTH: 47

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Primer Cm F

<400> SEQUENCE: 200

-continued

attttaaattct cgagtagagg atcccaacaa acgaaaattg gataaag 47

<210> SEQ ID NO 201
<211> LENGTH: 29
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer Cm R

<400> SEQUENCE: 201

acgcgttatt ataaaagcca gtcattagg 29

<210> SEQ ID NO 202
<211> LENGTH: 62
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer P11 F

<400> SEQUENCE: 202

tcgagagcgc tatagttggt gacagaatgg acatactatg atatattggt gctatagcgc 60

cc 62

<210> SEQ ID NO 203
<211> LENGTH: 58
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer P11 R

<400> SEQUENCE: 203

gggcgctata gcaacaatat atcatagtat gtccattctg tcaacaacta tagcgctc 58

<210> SEQ ID NO 204
<211> LENGTH: 38
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer PldhL F

<400> SEQUENCE: 204

gagctcgtcg acaaaccaac attatgacgt gtctgggc 38

<210> SEQ ID NO 205
<211> LENGTH: 30
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer PldhL R

<400> SEQUENCE: 205

ggatcctacc atgtttgtgc aaaataagtg 30

<210> SEQ ID NO 206
<211> LENGTH: 34
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Primer F-PnisA (EcoRV)

<400> SEQUENCE: 206

ttcagtgata tcgacatact tgaatgacct agtc 34

<210> SEQ ID NO 207

-continued

<211> LENGTH: 48
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Primer R-PnisA(PmeI BamHI)

<400> SEQUENCE: 207

ttgattagtt taaactgtag gatcctttga gtgcctcctt ataattta 48

<210> SEQ ID NO 208
 <211> LENGTH: 2460
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Template DNA

<400> SEQUENCE: 208

tgatccaaag gagggtagagg aaatggcgat gtttacgacc accgcaaaag ttattcagcc 60
 gaaaattcgt ggtttttatgt gcaccaccac ccacccgatt ggttgcgaaa aacgtgttca 120
 ggaagaaatc gcatacgcac gcgcgcaccc gccgaccagc ccgggtccga aacgtgtgct 180
 ggttattggc tgcagtacgg gctatggcct gagcaccctg atcaccgcgg cctttggtta 240
 tcaggccgca accctggggc tgtttctggc aggcgccgag accaaaggcc gtccggccgc 300
 ggccgggttg tataatacgg ttgcgttcga aaaagccgcc ctggaagcag gtctgtatgc 360
 acgttctctg aatggtagat cggttcgatc taccacgaaa gccgcacccg tggaagcaat 420
 taaacgtgat ctgggtaccg ttgatctggt ggtgtatagc attgcagcgc cgaaacgtac 480
 cgatccggcc accggcgtgc tgcataaagc gtgcctgaaa ccgattgggt caacctacac 540
 caatcgtagc gtgaacaccg ataagcaga agttaccgat gtgagtattg aaccggccag 600
 tccggaagaa atcgagata ccgtgaaagt tatgggtggc gaagattggg aactgtggat 660
 tcaggcactg agcgaagccg gcgtgctggc cgaaggcgca aaaaccgttg cgtattctta 720
 tattggcccg gaaatgacgt ggccgggtgta ttggagtggc accattggcg aagccaaaaa 780
 agatgttgaa aaagcggcga aacgcacac ccagcagtag ggctgtccgg cgtatccggt 840
 tgttgccaaa gcgctggtga ccagggccag tagcgccatt ccggtggtgc cgctgtatat 900
 ttgctgctg tatcgtgtta tgaagaaaa aggcacccat gaaggctgca ttgaacagat 960
 ggtgctgctg ctgacgacga aactgtatcc ggaaaatggt gcgccgatcg tggatgaagc 1020
 gggccgtgtg cgtgttgatg attgggaaat ggcagaagat gttcagcagg cagttaaaga 1080
 tctgtggagc caggtgagta cggccaatct gaaagatatt agcgattttg cagggttatca 1140
 gaccgaattt ctgcgtctgt ttggctttgg tattgatggt gtggattacg atcagccggt 1200
 tgatgttgaa gcggatctgc cgagcgccgc ccagcagtaa gtcaacaaag gaggggttaa 1260
 aatggttgat ttcgaatatt caataccaac tagaattttt ttcggtaaaag ataagataaa 1320
 tgtacttgga agagagctta aaaaatatgg ttctaaagtg cttatagttt atggtggagg 1380
 aagtataaag agaaatggaa tatatgataa agctgtaagt atacttgaaa aaaacagtat 1440
 taaattttat gaacttgtag gagtagagcc aaatccaaga gtaactacag ttgaaaaagg 1500
 agttaaataa tgtagagaaa atggagttga agtagtacta gctatagggt gaggaagtgc 1560
 aatagattgc gcaaagggtta tagcagcagc atgtgaatat gatggaaatc catgggatat 1620
 tgtgttagat gggtcaaaaa taaaaagggt gcttcctata gctagtatat taaccattgc 1680
 tgcaacagga tcagaaatgg atacgtgggc agtaataaat aatatggata caaacgaaaa 1740
 actaattgag gcacatccag atatggctcc taagttttct atattagatc caacgtatac 1800

-continued

```

gtataccgta cctaccaatc aaacagcagc aggaacagct gatattatga gtcatatatt 1860
tgaggtgtat tttagtaata caaaaacagc atatttgcag gatagaatgg cagaagcgtt 1920
attaagaact tgtattaaat atggaggaat agctcttgag aagccggatg attatgaggc 1980
aagagccaat ctaatgtggg cttcaagtct tgcgataaat ggacttttaa catatggtaa 2040
agacactaat tggagtgtac acttaatgga acatgaatta agtgcttatt acgacataac 2100
acacggcgta gggcttgcaa ttttaacacc taattggatg gagtatatat taaataatga 2160
tacagtgtac aagtttgttg aatatggtgt aaatgtttgg ggaatagaca aagaaaaaaa 2220
tcactatgac atagcacatc aagcaataca aaaaacaaga gattactttg taaatgtact 2280
aggtttacca tctagactga gagatgttgg aattgaagaa gaaaaattgg acataatggc 2340
aaaggaatca gtaaagctta caggaggaac cataggaaac ctaagaccag taaacgcctc 2400
cgaagtccta caaatattca aaaaatctgt gtaaacctac gtttaaactt acgcgtatga 2460

```

What is claimed is:

1. A recombinant microbial host cell comprising heterologous DNA molecules encoding polypeptides that catalyze substrate to product conversions for each step below:

- a) acetyl-CoA to acetoacetyl-CoA;
- b) acetoacetyl-CoA to 3-hydroxybutyryl-CoA;
- c) 3-hydroxybutyryl-CoA to crotonyl-CoA;
- d) crotonyl-CoA to butyryl-CoA; and
- e) butyryl-CoA to butyraldehyde,

wherein said microbial host cell produces 1-butanol by the action of an endogenous alcohol dehydrogenase, wherein said microbial host cell produces 1-butanol through the substrate to product conversions of (a) to (e) under aerobic conditions.

2. A host cell according to claim 1 wherein the polypeptide that catalyzes a substrate to product conversion of acetyl-CoA to acetoacetyl-CoA is acetyl-CoA acetyltransferase.

3. A host cell according to claim 1 wherein the polypeptide that catalyzes a substrate to product conversion of acetoacetyl-CoA to 3-hydroxybutyryl-CoA is 3-hydroxybutyryl-CoA dehydrogenase.

4. A host cell according to claim 1 wherein the polypeptide that catalyzes a substrate to product conversion of 3-hydroxybutyryl-CoA to crotonyl-CoA is crotonase.

5. A host cell according to claim 1 wherein the polypeptide that catalyzes a substrate to product conversion of crotonyl-CoA to butyryl-CoA is butyryl-CoA dehydrogenase.

6. A host cell according to claim 1 wherein the polypeptide that catalyzes a substrate to product conversion of butyryl-CoA to butyraldehyde is butyraldehyde dehydrogenase.

7. A host cell according to claim 1 wherein the cell is selected from the group consisting of: a bacterium, a *cyanobacterium*, a filamentous fungus and a yeast.

8. A host cell according to claim 7 wherein the cell is a member of a genus selected from the group consisting of *Clostridium*, *Zymomonas*, *Escherichia*, *Salmonella*, *Rhodococcus*, *Pseudomonas*, *Bacillus*, *Lactobacillus*, *Enterococcus*, *Alcaligenes*, *Klebsiella*, *Paenibacillus*, *Arthrobacter*, *Corynebacterium*, *Brevibacterium*, *Pichia*, *Candida*, *Hansenula* and *Saccharomyces*.

9. A host cell according to claim 8 wherein the cell is *Escherichia coli*.

10. A host cell according to claim 8 wherein the cell is *Alcaligenes eutrophus*.

11. A host cell according to claim 8 wherein the cell is *Bacillus licheniformis*.

12. A host cell according to claim 8 wherein the cell is *Paenibacillus macerans*.

13. A host cell according to claim 8 wherein the cell is *Rhodococcus erythropolis*.

14. A host cell according to claim 8 wherein the cell is *Pseudomonas putida*.

15. A host cell according to claim 8 wherein the cell is *Bacillus subtilis*.

16. A host cell according to claim 8 wherein the cell is *Lactobacillus plantarum*.

17. A host cell according to claim 8 wherein the cell is selected from the group consisting of *Enterococcus faecium*, *Enterococcus gallinarum*, and *Enterococcus faecalis*.

18. A host cell according to claim 8 wherein the cell is *Saccharomyces cerevisiae*.

19. A host cell according to claim 1 wherein the host cell is a facultative anaerobe.

20. A method for the production of 1-butanol comprising:

- i) providing the recombinant microbial host cell of claim 1; and
- ii) contacting the host cell of (i) with a fermentable carbon substrate under conditions whereby 1-butanol is produced.

21. A method according to claim 20 wherein the fermentable carbon substrate is selected from the group consisting of monosaccharides, oligosaccharides, and polysaccharides.

22. A method according to claim 20 wherein the carbon substrate is selected from the group consisting of glucose, sucrose, and fructose.

23. A method according to claim 20 wherein the conditions whereby 1-butanol is produced are anaerobic.

24. A method according to claim 20 wherein the conditions whereby 1-butanol is produced are microaerobic.

25. A method according to claim 20 wherein the host cell is contacted with the carbon substrate in minimal media.

26. A method according to claim 20 wherein the polypeptide that catalyzes a substrate to product conversion of acetyl-CoA to acetoacetyl-CoA is acetyl-CoA acetyltransferase.

27. A method according to claim 20 wherein the polypeptide that catalyzes a substrate to product conversion of acetoacetyl-CoA to 3-hydroxybutyryl-CoA is 3-hydroxybutyryl-CoA dehydrogenase.

209

28. A method according to claim 20 wherein the polypeptide that catalyzes a substrate to product conversion of 3-hydroxybutyryl-CoA to crotonyl-CoA is crotonase.

29. A method according to claim 20 wherein the polypeptide that catalyzes a substrate to product conversion of crotonyl-CoA to butyryl-CoA is butyryl-CoA dehydrogenase.

30. A method according to claim 20 wherein the polypeptide that catalyzes a substrate to product conversion of butyryl-CoA to butyraldehyde is butyraldehyde dehydrogenase.

31. A method according to claim 20 wherein the host cell is selected from the group consisting of: a bacterium, a *cyano-bacterium*, a filamentous fungus and a yeast.

32. A method according to claim 31 wherein the host cell is a member of a genus selected from the group consisting of *Clostridium*, *Zymomonas*, *Escherichia*, *Salmonella*, *Rhodococcus*, *Pseudomonas*, *Bacillus*, *Lactobacillus*, *Enterococcus*, *Alcaligenes*, *Klebsiella*, *Paenibacillus*, *Arthrobacter*, *Corynebacterium*, *Brevibacterium*, *Pichia*, *Candida*, *Hansenula* and *Saccharomyces*.

33. A method according to claim 32 wherein the host cell is *Escherichia coli*.

34. A method according to claim 32 wherein the host cell is *Alcaligenes eutrophus*.

35. A method according to claim 32 wherein the host cell is *Bacillus licheniformis*.

36. A method according to claim 32 wherein the host cell is *Paenibacillus macerans*.

37. A method according to claim 32 wherein the host cell is *Rhodococcus erythropolis*.

38. A method according to claim 32 wherein the host cell is *Pseudomonas putida*.

39. A method according to claim 32 wherein the host cell is *Bacillus subtilis*.

40. A method according to claim 32 wherein the host cell is *Lactobacillus plantarum*.

41. A method according to claim 32 wherein the host cell is selected from the group consisting of *Enterococcus faecium*, *Enterococcus gallinarum*, and *Enterococcus faecalis*.

42. A method according to claim 32 wherein the host cell is *Saccharomyces cerevisiae*.

43. A method according to claim 26 wherein the acetyl-CoA acetyltransferase has an amino acid sequence selected from the group consisting of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:129, SEQ ID NO:131, and SEQ ID NO:133.

210

44. A method according to claim 27 wherein the 3-hydroxybutyryl-CoA dehydrogenase has an amino acid sequence selected from the group consisting of SEQ ID NO:6, SEQ ID NO:135, SEQ ID NO:137, and SEQ ID NO:139.

45. A method according to claim 28 wherein the crotonase has an amino acid sequence selected from the group consisting of SEQ ID NO:8, SEQ ID NO:141, SEQ ID NO:143, and SEQ ID NO:145.

46. A method according to claim 29 wherein the butyryl-CoA dehydrogenase has an amino acid sequence selected from the group consisting of SEQ ID NO:10, SEQ ID NO:147, SEQ ID NO:149, SEQ ID NO:151, and SEQ ID NO:187.

47. A method according to claim 30 wherein the butyraldehyde dehydrogenase has an amino acid sequence selected from the group consisting of SEQ ID NO:12, SEQ ID NO:153, and SEQ ID NO:189.

48. A method according to claim 20 wherein the host cell is a facultative anaerobe.

49. A host cell according to claim 1 further comprising a heterologous DNA molecule that encodes a polypeptide that catalyzes the substrate to product conversion of butyraldehyde to 1-butanol.

50. The method according to claim 20 wherein the host cell further comprises a heterologous DNA molecule that encodes a polypeptide that catalyzes the substrate to product conversion of butyraldehyde to 1-butanol.

51. A host cell according to claim 1 wherein the host cell produces more butanol than a control cell lacking the heterologous DNA molecules encoding polypeptides that catalyze the substrate to product conversions of steps (a) to (e).

52. A host cell according to claim 50 wherein the polypeptide that catalyzes the substrate to product conversion of butyraldehyde to 1-butanol is butanol dehydrogenase.

53. A method according to claim 50 wherein the polypeptide that catalyzes the substrate to product conversion of butyraldehyde to 1-butanol is butanol dehydrogenase.

54. A method according to claim 53 wherein the butanol dehydrogenase has an amino acid sequence selected from the group consisting of SEQ ID NO:14, SEQ ID NO:16, SEQ ID NO:153, SEQ ID NO:155, and SEQ ID NO:157.

* * * * *